UNCLASSIFIED

AD NUMBER

AD480109

NEW LIMITATION CHANGE

TO

Approved for public release, distribution unlimited

FROM

Distribution authorized to U.S. Gov't. agencies and their contractors; Administrative/Operational Use; AUG 1964. Other requests shall be referred to Director of Security, Army Training and Doctrine Command, Fort Monroe, VA.

AUTHORITY

TRADOC, per DTIC Form 55, 10 Feb 1997

THIS PAGE IS UNCLASSIFIED

EMBASSY REQUEST

DEFENSE TECHNICAL INFORMATION CENTER
REQUEST FOR RELEASE OF LIMITED DOCUMENT
NOTE: This form may be classified if necessary. See instructions on reverse.

DTIC CONTROL NO. F7029001

USER ROUTING

SECTION I REQUESTING ORGANIZATION

REQUESTING ORGANIZATION AND ADDRESS
AUSTRALIAN EMBASSY
Publications Dept.
1601 MASS. AVE. NW RM 700
WASH. DC 20036
Attn: ROXANNE PARKER
DP 3037/54

DTIC USER CODE NO. 20011

DATE OF REQUEST

TYPE COPY AND QUANTITY
☒ Paper Copy 1 Copy(s) ☐ Microform ____ Copy(s)

METHOD OF PAYMENT (X ONE)
☒ Charge to NTIS Deposit Account No. 72714-9
☐ Bill My Organization to the Attention of. ____

SIGNATURE, TITLE, PHONE NUMBER OF REQUESTING OFFICIAL
Roxanne Parker 202-797-3045

SECTION II BIBLIOGRAPHIC INFORMATION

AD NUMBER 480 109

BIBLIOGRAPHY (Title, Report Number, Author(s))
VOL II. Historical data and Lanchester's theory of combat Pt. II. HELMBOLD, R.L. AUG 1964
REF NO. U433 (Ft. Belvoir or Ft. Monroe, VA)
CORG-SP-190

SECTION III REQUESTER JUSTIFICATION

REQUIRED FOR (Explain need in detail)

SECTION IV RELEASING AGENCY DECISION

RELEASING AGENCY (Use Post Office Address Format)
Army Training & Doctrine CMD
ATTN: ATIS-CIN
Ft Monroe, VA 23651

RELEASING AGENCY DECISION
☒ APPROVED FOR RELEASE TO THE ABOVE REQUESTER DOCUMENT TO BE PROVIDED BY RELEASING AUTHORITY.
☐ SANITIZED - PART OF TEXT REMOVED. DOCUMENT TO BE PROVIDED BY RELEASING AUTHORITY
☐ SANITIZED - DISTRIBUTION LIST AND BIBLIOGRAPHY REMOVED. DOCUMENT TO BE PROVIDED BY RELEASING AUTHORITY.
☐ DISAPPROVED. REASON FOR DISAPPROVAL
☐ ____
RELEASED UNDER EXCHANGE AGREEMENT

Distribution caveat should be changed from limited to "APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED"

TYPED NAME AND TITLE OF RELEASING OFFICIAL	TELEPHONE NO.	SIGNATURE	DATE
THOMAS M. McNEAR, LTC, MI Director of Security	(757) 727-2315	[signature]	10 FEB 97

DTIC FORM APR 95 55B

PREVIOUS EDITIONS MAY BE USED UNTIL STOCK IS EXHAUSTED

24 Flds

C-2 to A1

tech/ops CORG

COMBAT OPERATIONS RESEARCH GROUP

CORG-SP-190
August 1964

CORG STAFF PAPER
CORG-SP-190

HISTORICAL DATA AND LANCHESTER'S THEORY OF COMBAT

PART II

[illegible] THE ARMY CONTRACT
[illegible] COMMAND
[illegible]ERATIONS, INC.

tech ops CORG

COMBAT OPERATIONS RESEARCH GROUP

POST OFFICE BOX 116
FORT BELVOIR, VIRGINIA
22060

April 12, 1966

Headquarters
Defense Documentation Center
Defense Supply Agency
Cameron Station
Alexandria, Virginia 22314

Dear Sir:

Inclosed are 13 copies of CORG-SP-190; with/Change No. 1, 15 Apr 1965;

Subject: Historical Data and Lanchester's Theory of Combat Part II (U)

Date: August 1964
Classification: UNCLASSIFIED

Due to numerous requests and our depleting stock, it is requested that these 13 copies of CORG-SP-190 be included in your DDC technical and scientific document collection.

All distribution of this report is controlled. CORG Staff Papers are CORG "In-House" papers and all qualified DDC users shall request this document from Technical Operations Research, Combat Operations Research Group, ATTN: Security Officer, Post Office Box 116, Fort Belvoir, Virginia 22060.

Reproduction of this document in whole or in part is prohibited except with permission of Technical Operations Research, Combat Operations Research Group, Post Office Box 116, Fort Belvoir, Virginia 22060. However, the Defense Documentation Center is authorized to reproduce this document for United States Government purposes.

AR 70-31 does not apply.

The Combat Operations Research Group is a contract facility collocated with Headquarters, U. S. Army Combat Developments Command, Fort Belvoir, Virginia under DA Contract: DA-04-200-AMC-1623(X) (DDC User Code: 05318).

Sincerely yours,

JOSEPH T. [illegible]
Security Officer
CORG

1 Incl
as

CORG-SP-190
C1

COMBAT OPERATIONS RESEARCH GROUP
HEADQUARTERS
UNITED STATES ARMY COMBAT DEVELOPMENTS COMMAND
FORT BELVOIR, VIRGINIA

Changes No. 1
CORG STAFF PAPER
CORG-SP-190

15 April 1965

CORG STAFF PAPER

CORG-SP-190, Historical Data and
Lanchester's Theory of Combat, Part II

CORG-SP-190, August 1964 is changed as follows:

Make the following pen and ink changes:

1. Page 5, Battle 83, column headed "Initial attacker strength, x_0," "10,000" should read "10,269."

2. Page 6, second line from bottom, "A/S" should read "A/D."

3. Page 10, line 7, "$(c = C_x + C_y)$," should read "$(C = C_x + C_y)$."

4. Page 10, between second and third paragraphs insert the following paragraph:

> "In terms of F, the total casualty fraction $(F = C/X)$, the smallest of the 81 modified battles is Lexington with F = 0.002, and the largest is Calafat with F = 0.458. The average total casualty fraction (average of F's over 81 modified battles) is about 0.159."

5. Page 11, line 12, "correspondence" should read "correspondences."

6. Page 11, line 16, "a total casualty" should read "total casualty."

7. Page 11, line 21, ". . . 22, 43, 47, 74, 77, . . ." should read ". . . 22, 43, 45, 47, 74, 77, . . ."

8. Page 11, line 21, "(75 battles . . ." should read "(15 battles . . ."

9. Page 13, Battle 26, column headed "ϵ", "0.137" should read "0.139."

10. Page 14, Battle 56, column headed "ϵ", "1.000" should read "0.065."

11. Page 14, Battle 56, column headed "λ", "0.065" should read "0.032."

12. Page 14, Battle 75, column headed, "μ", "1.1092" should read "1.092."

13. Page 17, line 4, "$\lambda = \epsilon \mid t$" should read "$\lambda = \epsilon / t$."

14. Page 25, second line from bottom, "ln μ", should read "ln ϵ."

15. Page 26, line 1, "ln μ", should read "ln ϵ."

16. Page 29, line 6, "ln ϵ (modified data) are . . ." should read ". . . ln ϵ is normally distributed. Estimated mean and standard deviation of ln ϵ (modified data) are - 2.034 and 0.757, respectively."

17. Page 31, line 3, "Attacking Side Victorious (49 Battles)," should read "Attacking Side Victorious (47 Battles)."

18. Page 31, line 7, both columns headed "Surviving Fraction of Defender, d" "90" should read "%".

19. Page 33, column headed "Level 3", "($1.500 \leq x_0 / y_0 < 00$)" should read "($1.500 \leq x_0 / y_0 < \infty$)."

20. Page 34, line 5, "ratios of 2,000 or greater" should read "ratios of 2.000 or greater."

21. Page 34, line 8, "ratios of 3,000 or greater" should read "ratios of 3.000 or greater."

22. Page 34, line 10, "ratio of 5,000 or greater." should read "ratio of 5.000 or greater."

23. Page 34, line 12, "ratio of 10,000 or greater." should read "ratio of 10.000 or greater."

24. Page 37, line 3, "diagram at μ" should read "diagram of μ."

25. Page 37, line 6 from bottom, "Residual Advantage = ln - b - c ln (x_0 / y_0)" should read "Residual Advantage = ln μ - b - c ln (x_0 / y_0)."

26. Page 39, line 2, ". . . FORCE RATIO, x_0 / y_0" should read ". . . FORCE RATIO, ln x_0 / y_0."

27. Page 43, Footnote 17, ". . . calculated the difference . . ." should read ". . . calculated from the difference . . ."

28. Page 47, line 7, ". . . an advantage . . ." should read ". . . on advantage . . ."

29. Page 65, line 6, "$c_{\ln \epsilon \mid \ln \mu} = 0.889$" should read "$s_{\ln \epsilon \mid \ln t}$ 0.880."

30. Page 71, Figure caption, " . . . c, against total force, x, for phase - 2 data." should read " . . . C, against total force, X, for phase - 2 data."

31. Page 71, curve label, "c = 0.15x" should read "C = 0.15X."

32. Page 75, Footnote 26, " . . . on the defindina side . . ." should read " on the defending side . . ."

33. Page 83, Footnote, "by equating deserved fraction . . ." should read "by equating observed fraction . . ."

34. Page 86, line 6, " . . . does the phase - 1." should read " . . . does the phase - 1 data."

35. Page 86, line 13, " . . . 16, 18, 22, 43, . . ." should read " . . . 16, 18, 21, 22, 43, . . ."

36. Page 88, line 7 from the bottom, " . . . a large F-ratio value . . ." should read " . . . a larger F-ratio value."

37. Page 89, line 8 from the bottom, " . . . for case of" should read " . . . for the case of . . ."

38. Page 92, line 8, "rigorous exact . . ." should read "rigorous, exact . . ."

39. Page 92, line 5 from the bottom, " . . . intervals and region . . ." should read " . . . intervals and regions . . ."

40. Page 92, lines 1, 2, & 3 from bottom, " . . . on the maximum likelihood . . ." should read " . . . on the maximum likelihood estimates of the regression parameters. Note that these maximum likelihood estimates plot" and continue with "as points on . . ."

41. Page 93, line 8, " . . . does not include . . ." should read " . . . does include . . ."

42. Page 96, line 11, "phase - 2 Harbottle, and Non-Harbottle . . ." should read "phase - 2, Harbottle, and Non-Harbottle . . ."

43. Page 96, line 1 from bottom. " . . . with phase - 2 data." should read " . . . with phase - 1 data."

44. Page 98, line 6 from bottom, "poorer agreement with a small proportion . . ." should read " . . . poorer agreement with phase - 1 data than those with a small proportion . . ."

45. Page 107, line 10, " . . . C = 15X . . ." should read " . . . C = 0.15X . . ."

46. Page 110, Footnote 40, "See footnote 37." should read "See footnote 39."

47. Page 111, Footnote 42, "See footnote 39." should read "See footnote 41."

48. Page 111, line 3, "ln μ = 0.1 ln x_0 / y_0" should read "ln μ = 0.1 - 0.3 ln (x_0 / y_0)."

49. Page 113, line 12 from bottom, "toward bitterness values" should read "toward lower bitterness values."

50. Page 113, Footnote 44, "See footnote 41." should read "See footnote 43."

51. Page 113, line 2 from bottom, "Data in accuracies large . . ." should read "Data inaccuracies are large . . ."

52. Page 114, Footnote 45, "See footnote 39." should read "See footnote 41."

53. Page 124, line 14 from bottom, "Residual Advantage = 0.845." should read "Residual Advantage = -0.845."

54. Page 132, Reference 8, "Furman" should read "Forman."

55. Page 13, Battle 43, "Kerygaom" should read "Korygaom."

tech/ops CORG

COMBAT OPERATIONS RESEARCH GROUP

CORG STAFF PAPER

CORG-SP-190

HISTORICAL DATA AND LANCHESTER'S THEORY OF COMBAT

PART II

by

Robert L. Helmbold

This work was sponsored by Headquarters, United States Army Combat Developments Command, Fort Belvoir, Virginia under Department of the Army Contract No. DA-19-020-AMC-00318X.

August 1964

This paper represents the individual views of the author and does not necessarily represent the opinion or approval of the Combat Operations Research Group, the United States Combat Developments Command or of the Department of the Army.

ABSTRACT

The purpose of this paper is to test the validity of the conclusions reached in a previous work (CORG-SP-128, Historical Data and Lanchester's Theory of Combat). Data on 83 historical battles are studied for information bearing on this problem. The information obtained from these battles is compared with those previously obtained, and the results of the comparison are carefully analyzed. With a few exceptions, the validity of the earlier findings is confirmed.

FOREWORD

Quote from " The Annual Report of the Chief of Staff, U.S. Army, 1935" , as presented in Maneuver in War, by Charles A. Willoughby, Military Service Publishing Co., Harrisburg, Pa.

" More than most professions, the military is forced to depend upon intelligent interpretation of the past for signposts charting the future. Devoid of opportunity, in peace, for self-instruction through actual practice of his profession, the soldier makes maximum use of historical record in assuring the readiness of himself and his command to function efficiently in emergency.

" The facts derived from historical analysis, he applies to conditions of the present and the proximate future, thus developing a synthesis of approximate method, organization,and doctrine.

" But the military student does not seek to learn from history the minutiae of method and techniques. In every age, these are decisively influenced by the characteristics of weapons currently available and by the means at hand, for maneuvering, supplying,and controlling combat forces. But research does bring to light those fundamental principles and their combination and application which, in the past, have been productive of success.

" These principles know no limitation of time. Consequently, the Army extends its analytical interest to the dust-buried accounts of wars long past as well as to those still reeking with the scent of battle. It is the object of the search that dictates the field for its pursuit. Those callow critics who hold that only in the most recent battles are there to be found truths applicable to our present problems have failed utterly to see this.

" They apparently cling to the fatuous hope that in historical study is to be found a complete digest of the science of war rather than simply the basic and inviolable laws of the art of war .... " .

CONTENTS

CONTENTS (Continued)

CONTENTS (Continued)

CONTENTS (Continued)

CONTENTS (Continued)

Page

FIGURES (Continued)

CONTENTS (Concluded)

HISTORICAL DATA AND LANCHESTER'S THEORY OF COMBAT

PART II

INTRODUCTION

Purpose

The purpose of this study was to test the findings obtained in Reference 1 against additional historical data.

Background

Reference 1 presented an analysis of 92 historical battles based on certain parameters suggested by Lanchester's square-law model of combat. The reader is referred to Reference 1 for a discussion of background and scope, theoretical and methodological principles, data sources and errors, presentations and analysis of preliminary results, and suggestions for possible applications.[1] It is desirable to check the preliminary findings of Reference 1 against additional historical battle data.

Scope

Data from 83 historical battles is presented in this paper, analyzed as in Reference 1, and compared with the results obtained in Reference 1.

LIMITATIONS OF THE METHOD

As in Reference 1, the quality of the data sets limits on the results that can be obtained. Most of the comments of Reference 1 relevant to this topic

---

[1]In preparing this paper, the author has assumed that the reader will be familiar with Reference 1 and will be able to review it while studying the following material.

are appropriate here. Additional remarks pertinent to the historical data used in this study are given below.

Identification of Attacker and Defender

The identifications given in Table I are based either on the identifications given in the sources consulted or on the author's impression, based on the narrative accounts, of which side behaved most aggressively. When allied nations participated on either side, an attempt is made to assign identifications which correspond either to the nationality of the largest body of troops or to the nationality of the commander of the allied forces. For some of the battles, especially those prior to the year 1800[2], the names of participating nations are altered to reflect the modern nationality of the region. For example, Reference 2 gives the combatants at the battle of Heraclea as Epirots and Romans. Table I gives the combatants as Greeks and Romans.

Duration of Engagement

In Reference 1, the duration of some of the older battles was adjusted on the assumption that there was no night fighting. No such adjustment was attempted for the data used in this study. Thus, the duration data given in Table I represents an estimate of the total elapsed time from initiation to termination of the battle.[3]

---

[2] All dates are to be read as A.D. unless specifically identified as B.C.

[3] See Reference 1 for additional discussion of data limitations, especially with regard to initial strength and casualty estimates.

TABLE I

BATTLE DATA

[illegible]	[illegible]	Battle	Source	War, Campaign, or Article	Attacker	Defender	Initial attacker strength, x_0	Initial defender strength, y_0	Attacker's casualties, C_x	Defender's casualties, C_y	Duration of engagement, T, hr	Nominal victor
[illegible]	[illegible]	Heraclea	H	Pyrrhus' Invasion of Italy	Greek	Roman	30 000	35 000	4 000	7 000		A
[illegible]	[illegible]	Asculum	H	Pyrrhus' Invasion of Italy	Roman	Greek	45 000	45 000	6 000	3 000		D
[illegible]	[illegible]	Mursa	H	Revolt of Magnentius	Eastern Roman	Western Roman	80 000	100 000	30 000	24 000		A
[illegible]	[illegible]	Indus	H	Mongol Invasion of Kharismia	Mongol	Indian	300 000	30 000	[illegible]	19 000		A
[illegible]	[illegible]	Sempach	H	War of Sempach	Austrian	Swiss	6 000	1 500	[illegible]	150		D
[illegible]	[illegible]	Ravenna	W	War of the Holy League	French	Spanish	20 000	12 000	[illegible]	9 000	7.0	A
[illegible]	[illegible]	Marignano	W	Italian Wars	Swiss	French	30 000	46 000	5 000	4 000	7.0	D
[illegible]	[illegible]	Dreux	F B	Dreux	British	French	16 000	12 000	3 000	3 000		A
[illegible]	[illegible]	Leipzig	H	Thirty Year's War	German	Swede	44 000	40,300	[illegible]	2 700		D
[illegible]	[illegible]	[illegible]	H	Thirty Year's War	Swede	German	40 000	30 000	[illegible]	2 070		D
[illegible]	[illegible]	[illegible]	W	Thirty Year's War	Swede	German	15 000	12 000	5 000	5 000	[illegible]	A
[illegible]	[illegible]	[illegible]	F B	Thirty Year's War	German	Swede	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	D
[illegible]	[illegible]	[illegible]	H	Wars of Louis XIV	French	German	17 000	[illegible]	[illegible]	6.000		[illegible]
[illegible]	[illegible]	[illegible]	F B.	[illegible]	French	German	[illegible]	37 000	[illegible]	5 000	10.0	A
[illegible]	[illegible]	[illegible]	H	Spanish Succession	British	French	[illegible]	60 000	[illegible]	40 000		A
[illegible]	[illegible]	[illegible]	H	[illegible]	Russian	Swede	40 000	[illegible]	[illegible]	10 000	[illegible]	A
[illegible]	[illegible]	[illegible]	H	War of Austrian Succession	Prussian	Austrian	[illegible]	[illegible]	[illegible]	6 000		A
[illegible]	[illegible]	[illegible]	H	Seven Year's War	French	English	900	1,400	[illegible]	477	[illegible]	A
[illegible]	[illegible]	[illegible]	H	Seven Year's War	British	French	3 000	4 000	1 000	400		D
[illegible]	[illegible]	[illegible]	H	First Mysore War	Indian	British	[illegible]	11 000	[illegible]	[illegible]	[illegible]	D
[illegible]	[illegible]	[illegible]	H	First Mysore War	British	Indian	[illegible]	65 000	[illegible]	10 000	12.0	A
[illegible]	[illegible]	[illegible]	H	First Mysore War	British	Indian	10 000	90 000	[illegible]	[illegible]		A
[illegible]	[illegible]	[illegible]	F B	[illegible]	French	Austrian	45 000	[illegible]	[illegible]	7 000	[illegible]	A
[illegible]	[illegible]	[illegible]	F B	[illegible]	French	Austrian	43 000	21 000	[illegible]	[illegible]	[illegible]	A
[illegible]	[illegible]	[illegible]	F B	[illegible]	French	British	[illegible]	14 000	[illegible]	[illegible]	[illegible]	D
[illegible]	[illegible]	[illegible]	H	[illegible]	British	French	4 000	[illegible]	[illegible]	[illegible]		A
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	French	British	13 000	20 000	[illegible]	[illegible]		D
[illegible]	[illegible]	[illegible]	H	Campaign of Wagram	Austrian	French	90 000	[illegible]	[illegible]	[illegible]		D

TABLE I (Continued)

BATTLE DATA

No.	Date	Battle	Source	War, Campaign or Attack	Attacker	Defender	Initial attacker strength, x_0	Initial defender strength, y_0	Attacker's casualties C_x	Defender's casualties C_y	[illegible]	[illegible]
29	[illegible]	Corunna	H	Peninsular War	French	British	20,000	14,000	2,000[l]	900[l]		D
30	[illegible]	Eckmuhl	H	Wagram	French	Austrian	90,000	74,000	2,400	5,000		A
31	[illegible]	Bussaco	E B	Bussaco	French	British	40,000	51,000	5,000[l]	1,300[l]	9.[illegible]	D
32	[illegible]	Mohilev	H	Moscow Campaign	Russian	French	60,000	20,000	4,000[l]	4,000[l]		D
33	[illegible]	Malo Jaroslavetz	H	Moscow Campaign	French	Russian	15,000	24,000	5,000[l]	6,000[l]		D
34	[illegible]	Polotsk	H	Moscow Campaign	French	Russian	33,000	30,000	1,000	2,000[k]	2.1	A
35	[illegible]	Smolensk	H	Moscow Campaign	French	Russian	50,000	60,000	9,000[l]	10,000[l]		A
36	[illegible]	Valutinagora	H	Moscow Campaign	French	Russian	30,000	40,000	7,000[l]	7,000[l]		A
37	[illegible]	Sauroren	H	Peninsular War	French	British	25,000	12,000	3,000[l]	2,600[l]	12.9	D
38	[illegible]	Wartenburg	H	Campaign of Leipzig	Prussian	French	60,000	16,000	5,000[l]	500[l]	6.5	A
39	[illegible]	Brienne	H	Allied Invasion of France	German	French	30,000	14,000	4,000	3,000		D
40	[illegible]	Craonne	H	Allied Invasion of France	French	German	55,000	50,000	9,000	7,000		A
41	[illegible]	Rheims	H	Allied Invasion of France	French	Prussian	30,000	14,000	1,300[l]	6,000[l]		A
42	[illegible]	Toulouse	H	Peninsular War	British	French	24,000	36,000	4,650	3,000		A
43	[illegible]	Korygaum	H	Third Mahratta War	Indian	British	25,000	1,000	600[l]	275[l]	12.0	D
44	[illegible]	Warsaw	H	Second Polish Uprising	Russian	Pole	60,000	40,000	[illegible],500	9,000[k]		A
45	[illegible]	Alamo	E B	Alamo	Mexican	Texan	4,000	180	500[k]	175[k]	312.9	A
46	[illegible]	Maharajpur	H	Gwalior War	British	Indian	14,000	18,000	787	3,000		A
47	[illegible]	Meeanee	H	Scinde Campaign	Persian	British	20,000	2,400	5,000[l]	256		D
48	[illegible]	Sobraon	H	First Sikh War	British	Indian	15,000	2[illegible],000	2,383	8,000[l]		A
49	[illegible]	Boomplaats	H		British	Boer	1,050	1,000	40	14		A
50	[illegible]	[illegible]	H	Second Sikh War	[illegible]	[illegible]	11,000	4,000	300[l]	500[l]		A
51	[illegible]	Calafat	H	Crimean War	Russian	Turk	40,000	30,000	20,000[l]	12,000[l]		D
52	[illegible]	Lexington	E B	Lexington	Confederate	Federal	15,000	2,400	100	159	12.0	A
53	[illegible]	Wilson's Creek	H	American Civil War	Federal	Confederate	6,000	16,000	1,236	1,095		D
54	[illegible]	Cedar Mountain	H	American Civil War	Federal	Confederate	20,000	15,[illegible]	2,000	900	5.0	D
55	[illegible]	La Puebla	H	Franco-Mexican War	French	Mexican	7,500	12,000	456[l]	[illegible]		D
56	[illegible]	Pea Ridge	H	American Civil War	Federal	Confederate	[illegible]	16,000	1,060[l]	1,000[l]	60.0	A
57	[illegible]	Perryville	E B	Perryville	Federal	Confederate	[illegible]	16,000	4,211	3,000		A

TABLE I (Concluded)

BATTLE DATA

No.	Date	Battle	Source	War, Campaign or Rebellion	Attacker	Defender	Initial attacker strength x_0	Initial defender strength y_0	Attacker casualties	Defender casualties	Duration	Victor
58	[illegible]	Chancellorsville	H	American Civil War	Federal	Confederate	6,000	[illegible]	[illegible] l	[illegible] l		[illegible]
59	[illegible]	[illegible] Hill	E B	Shenandoah Valley	Confederate	Federal	6,000	2,500	[illegible] l	[illegible] l	[illegible]	[illegible]
60	[illegible]	Shiloh	L	American Civil War	Confederate	Federal	40,000	61,000	10,[illegible] m	11,[illegible] m	[illegible]	D
61	[illegible]	Gettysburg	L	American Civil War	Confederate	Federal	75,000	[illegible] 000	25,[illegible] m	[illegible] m	[illegible]	[illegible]
62	[illegible]	Vicksburg	L	American Civil War	Federal	Confederate	71,000	31,000	[illegible]	[illegible] m	[illegible]	A
63	[illegible]	Atlanta	L	American Civil War	Federal	Confederate	100,000	[illegible] 000	[illegible] m	[illegible] m	[illegible]	A
64	[illegible]	Cold Harbor	L	American Civil War	Federal	Confederate	10[illegible],000	[illegible] 000	[illegible] m	1,[illegible] m	[illegible]	[illegible]
65	[illegible]	Kennesaw Mt	H	American Civil War	Federal	Confederate	90,000	[illegible] 000	[illegible] l	[illegible] l		D
66	[illegible]	Nashville	I	American Civil War	Federal	Confederate	49,000	[illegible] 000	[illegible] m	1,[illegible] m	[illegible]	A
67	[illegible]	[illegible]	H	American Civil War	Federal	Confederate	6,000	[illegible] 000	1,[illegible] l	[illegible] l		D
68	[illegible]	Spotsylvania	L	American Civil War	Federal	Confederate	90,000	50,000	1[illegible] m	9,[illegible] m	[illegible]	A
69	[illegible]	Five Forks	L	American Civil War	Federal	Confederate	29,000	19,000	1,[illegible] m	[illegible] m	[illegible]	A
70	[illegible]	Custozza	H	Seven Weeks's War	Italian	Austrian	110,000	[illegible] 000	[illegible]	[illegible] w		D
71	[illegible]	[illegible]	H	Franco-German War	German	French	22,500	[illegible] 000	[illegible] w	1,000 w		D
72	[illegible]	St. Quentin	H	Franco-German War	French	German	40,000	[illegible] 000	1,[illegible] w	[illegible] w		D
73	[illegible]	[illegible]	H	Russo-Turkish War	Russian	Turk	20,000	15,000	1,[illegible] w	5,[illegible] k		[illegible]
74	[illegible]	Bronkhorst Spruit	H	First Boer War	Boer	British	150	259	[illegible] w	[illegible] w		A
75	[illegible]	Fish Creek	H	Riel's Second Rebellion	Canadian	Rebel	400	250	[illegible] w	[illegible] w		D
76	[illegible]	Slivnitza	H	Serbo-Bulgarian War	Serb	Bulgar	25,000	15,000	2,[illegible] l	3,[illegible] l	[illegible]	D
77	[illegible]	Atbara	H	Sudan Campaign	British	Sudan	14,000	1[illegible] 000	[illegible] k	[illegible] k		A
78	[illegible]	[illegible]	H	Russo-Japanese War	Japanese	Russian	40,000	10,000	[illegible] k	[illegible]		A
79	[illegible]	Telissu	E B	Russo-Japanese War	Japanese	Russian	35,000	[illegible] 000	1,[illegible] l	3,[illegible] l	[illegible]	A
80*	[illegible]	Attu	M	World War II	American	Japanese	11,000	[illegible] 000	1,[illegible] w	2,[illegible] k	[illegible]	A
81	[illegible]	Kwajalein South	M	World War II	American	Japanese	21,342	5,112	[illegible] m	[illegible] k	[illegible]	A
82	[illegible]	Kwajalein North	M	World War II	American	Japanese	30,104	3,563	[illegible] m	3,[illegible] k	[illegible]	A
83	[illegible]	Eniwetok	M	World War II	American	Japanese	10,000	3,431	216 m	3,[illegible] l	[illegible]	A

*[illegible] battle

Source		Victor		Casualties	
H	Harbottle's Dictionary	D	Defender	l	losses
E B	Encyclopaedia Britannica	A	Attacker	k	killed
[illegible]	Warfare			w	killed and wounded
L	Life Magazine			m	killed, wounded and missing
M	Morrison				

RESULTS OF THE ANALYSIS OF 83 BATTLES

Data

Data on the historical battles treated in this portion of the study were assembled from References 2, 3, 4, 5, and 6. The principal quantitative components of these data (together with the identifications of attacker, defender, and victor) are presented in Table I.

Population and Sample

Certain features of the sample battles listed in Table I which might be of assistance in judging the nature of the population are given below.

1. Sample Size

Table I contains data on a total of 83 distinct historical battles. These battles are also distinct from those forming the principal data of Reference 1. For each of these battles the name or designation of the battle; the year in which it occurred, the war, campaign, or article to which the data refer; the source of the data used, the identification of attacker and defender; the estimated initial strength on each side; the estimated casualties suffered by each side (together with a notation approximately identifying the casualty criterion involved) and an identification of the victor are listed. For 36 of these battles an estimate of the time duration is also supplied.

Note: Of the 83 battles listed in Table I, only 81 will be used in the remainder of this chapter. Battles numbered 21 and 22, (Sholingur and Porto Novo) are deleted because, firstly, the data for these battles deviate greatly from that of the other battles analyzed here and in Reference 1, secondly, because the data for these battles is such that the regression of activity ratio, A/S, and of defender advantage, ln μ, on logarithmic force ratio, ln x_0/y_0, is strongly affected if these

battles are included, and thirdly, because the narrative accounts of these battles suggest that the data on these battles is untrustworthy (these narratives are reproduced in App. A). For ease in referring to the various cases, the 81 battles will be referred to as the modified battle data and the 83 battles will be referred to as the phase-2 data.[4] The 92 battles studied in Reference 1 will be referred to as the phase-1 data.

Data for the Excluded Battles (Porto Novo and Sholingur) is included in the tables only where explicitly identified. Data points for these excluded battles are always identified by triangular symbols whenever such data is included in a scatter diagram.

2. Distribution of Modified Data in Time

Modified battle dates vary from 280 B. C. to 1944, and 17 of these battles occurred prior to 1750. Figure 1 shows the distribution of modified battles in time by time-intervals of varying length. As for the phase-1 data, the high degree of clustering tends to reflect the occurrence of certain periods of general military activity.

3. Distribution of Modified Data in Space

Of the 81 modified battles, 43 were fought in the Eurasian area, 21 in North America, 9 in East or Central Asia, 4 in Africa, and 4 on Pacific Ocean islands.

---

[4]See Appendix B for detailed description of the various battle groupings used.

Time Interval	Number of Battles
500 B.C. - 1 B.C.	XX
1 (A.D.) - 499	X
500 - 999	
1000 - 1499	XX
1500 - 1599	XXX
1600 - 1699	XXXXXX
1700 - 1739	XX
1740 - 1759	XX
1760 - 1779	X
1780 - 1799	XXX
1800 - 1819	XXXXXXXXXXXXXXXXXXX
1820 - 1839	XX
1840 - 1859	XXXXXX
1860 - 1879	XXXXXXXXXXXXXXXXXXXXXX
1880 - 1899	XXXX
1900 - 1919	XX
1920 - 1939	
1940 - 1960	XXXX

Figure 1. Distribution of battles in time for modified battle data.

4. Distribution of Modified Data Among Countries

The nations which participated in the modified battles, and the approximate frequency of participation in these battles[5], were:

France	31 Battles
United States of America	21 Battles
Britain (or England)	19 Battles
Confederate States of America	17 Battles
Russia	12 Battles
Prussia (or Germany)	10 Battles
India	8 Battles
Austria	7 Battles
Rome[6]	6 Battles
Japan	6 Battles
Sweden	5 Battles
Greece	2 Battles
Turkey	2 Battles
South Africa	2 Battles
Mexico	2 Battles
Switzerland	2 Battles
Canada	2 Battles
Eastern Roman Empire	1 Battle
Mongolia	1 Battle
Spain	1 Battle
Persia	1 Battle
Poland	1 Battle
Yugoslavia	1 Battle
Italy	1 Battle
Texas	1 Battle

[5] It is difficult to know how much weight to give to participation with allies under a united command. Additional difficulties arise from the change in name or government of a region.

[6] Includes Western Roman Empire and Holy Roman Empire.

5. Magnitude of the Modified Battles

In terms of X, the total number of troops involved ($X = x_y + y_0$), the smallest of the 81 modified battles is Bronkhurst Spruit with $X = 409$, and the largest is Indus with $X = 330,000$. The overall total forces (sum of X's for all 81 modified battles) amount to 5,241,940, and the average total force (average of X's over 81 battles) is about 64,715.

In terms of C, the total number of casualties ($c = C_X + C_y$), the smallest of the 81 modified battles is Fish Creek with $C = 79$, and the largest is Mursa with $C = 54,000$. The overall total casualties (the sum of C's for all 81 modified battles) amounts to 749,930, and the average total casualties (average of C's over 81 modified battles) is about 9,258.

In terms of t, the duration of battle, the shortest of the 35 modified battles for which this datum is recorded is Cold Harbor with $t = 1$ hour, and the longest is Vicksburg with $t = 1800$ hours (75 days). The overall total battle duration (the sum of t's for 35 modified battles amounts to 5,102.5 hours, and the average modified battle duration (average of t's over 35 modified battles) is about 145.8 hours.

In terms of M, the number of battle man-hours ($M = Xt$), the smallest of the 35 modified battles for which this datum is available is Monongaheia with $M = 6,900$, and the largest is Vicksburg with $M = 219,600,000$. The overall total modified battle man-hours (the sum of M's for 35 modified battles) amount to 470,570,060, and the average number of battle man-hours (average of M's over 35 modified battles) is about 13,444, 486.

6. Numerical Superiority

Of the 81 modified battles, those with the smallest force ratio (x_0/y_0) are Ste. Foy and Wilson's Creek, both having $x_0/y_0 = 0.375$, and the one with the largest force ratio is Korygaom with $x_0/y_0 = 25.000$. The average force ratio (over 81 modified battles) is about 2.349.

7. Victorious Side

In 47 of the 81 modified battles the attacker is given credit for the victory. In the remaining 34 battles the defender is counted as victorious.

Findings

Lanchester parameters for each of the phase-2 battles were estimated from the data of Table I using procedures described in Reference 1, and are tabulated in Table II.

Correspondence Between Parameters and Phenomena

In Reference 1, it was argued that the names given to the theoretical symbols appearing in Lanchester's theory provide a valid correspondence between the Lanchester parameters and real-world phenomena. It is important to determine whether these correspondence persist in the modified data.

1. Bitterness

Figure 2 shows a linear scatter diagram of bitterness, ϵ, against a total casualty fraction, F, for the phase-2 data, together with a graph of the function $\epsilon = e^F - 1$. The bulk of the phase-2 data appear to follow the curve $\epsilon = e^F - 1$ although more phase-2 data points exhibit sizable deviations from the curve than phase-1 data points. Phase-2 battles which exhibit a marked deviation from the trend of Figure 2 are [7]: Battles numbered 4, 5, 15, 16, 18, 21, 22, 43, 47, 74, 77, 80, 81, and 82 (75 battles in all). When these 15 exceptional battles are deleted from the 83 phase-2 battles, the remaining 68 battles will be termed the Censored battles [8].

As in Reference 1, we conclude that the Lanchester bitterness parameter, ϵ, reflects the intuitive concept of combat bloodiness,

---

[7] Numbers given refer to the designations of Table I.

[8] See Appendix B for detailed description of the various battle groupings used.

TABLE II

BATTLE DATA PARAMETERS

No.	Battle	a	d	μ	D/A	ϵ	λ	α	δ	A	D	x_0/y_0
1	Heraclea	0.867	0.800	0.831	0.508	0.179	-	-	-	-	-	0.857
2	Asculum	0.867	0.933	1.399	1.931	0.099	-	-	-	-	-	1.000
3	Mursa	0.625	0.760	1.201	0.923	0.359	-	-	-	-	-	0.800
4	Indus	0.933	0.367	0.386	14.891	0.254	-	-	-	-	-	10.000
5	Sempach	0.750	0.920	1.688	45.573	0.155	-	-	-	-	-	4.000
6	Ravenna	0.850	0.250	0.544	0.822	0.448	1.538	2.826	0.837	1.696	1.394	1.667
7	Marignano	0.833	0.913	1.355	0.781	0.129	0.441	0.326	0.598	0.499	0.390	0.652
8	Dreux	0.812	0.750	0.881	1.381	0.244	-	-	-	-	-	1.333
9	Leipzig	0.818	0.932	1.592	3.066	0.118	-	-	-	-	-	1.100
10	Alltendorf	0.942	0.948	1.053	1.108	0.056	-	-	-	-	-	1.000
11	Lutzen	0.667	0.583	0.918	1.316	0.467	1.121	1.222	1.029	0.977	1.286	1.250
12	Wittstock	0.633	0.773	1.219	2.764	0.343	0.034	0.028	0.042	0.021	0.057	1.364
13	Nordlingen	0.765	0.571	0.785	0.909	0.386	-	-	-	-	-	1.214
14	Fleurus	0.944	0.865	0.655	0.634	0.091	0.219	0.334	0.143	0.274	0.174	1.216
15	Blenheim	0.788	0.333	0.652	0.320	0.495	-	-	-	-	-	0.867
16	Lissno	0.750	0.333	0.702	3.500	0.548	0.110	0.156	0.077	0.059	0.205	2.667
17	Sohr	0.778	0.829	1.123	0.333	0.217	-	-	-	-	-	0.514
18	Monongahela	0.889	0.374	0.494	0.101	0.331	2.644	5.354	1.306	8.329	0.840	0.643
19	Ste. Foy	0.667	0.900	1.710	0.411	0.206	-	-	-	-	-	0.375
20	Pollicore	0.975	0.962	0.811	34.787	0.0314	0.094	0.116	0.076	0.016	0.556	7.273
21	Porto Novo[a]	0.964	0.846	0.499	0.004	0.078	0.156	0.313	0.078	2.397	0.010	0.131
22	Sholingur[a]	0.990	0.938	0.405	0.003	0.025	-	-	-	-	-	0.125
23	Jemappes	0.844	0.720	0.772	1.339	0.235	0.706	0.915	0.545	0.508	0.981	1.800
24	Wattignies	0.930	0.891	0.809	2.290	0.091	0.109	0.135	0.089	0.072	0.166	1.870
25	Alexandria	0.792	0.895	1.370	3.828	0.161	0.779	0.563	1.056	0.394	1.508	1.429

[a] Excluded battle

TABLE II (Continued)

BATTLE DATA PARAMETERS

No.	Battle	a	d	μ	D/A	ϵ	λ	α	δ	A	D	x_0/y_0
[illegible]	Rolica	0.900	0.833	0.789	1.105	0.137	-	-	-	-	-	1.333
[illegible]	Vimiero	0.769	0.962	2.355	2.344	0.100	-	-	-	-	-	0.650
[illegible]	Abensburg	0.912	0.967	1.598	2.017	0.056	-	-	-	-	-	0.889
[illegible]	Coruna	0.900	0.943	1.308	3.493	0.079	-	-	-	-	-	1.429
[illegible]	Eckmuhl	0.972	0.934	0.656	0.604	0.044	-	-	-	-	-	1.184
[illegible]	Bussaco	0.938	0.975	1.551	5.920	0.041	0.122	0.079	0.190	0.050	0.298	1.569
32	Mohilev	0.933	0.964	1.355	8.436	0.050	-	-	-	-	-	2.143
[illegible]	Malo-Jaroslowetz	0.667	0.750	1.127	0.496	0.341	-	-	-	-	-	0.625
[illegible]	Polotsk	0.970	0.900	0.560	0.380	0.057	0.683	1.219	0.383	1.108	0.421	1.100
[illegible]	Smolensk	0.820	0.833	1.035	0.745	0.190	-	-	-	-	-	0.833
[illegible]	Valutinagora	0.767	0.825	1.136	0.726	0.226	-	-	-	-	-	0.750
[illegible]	Sauroren	0.980	0.783	0.764	2.534	0.177	0.353	0.462	0.270	0.222	0.562	2.083
[illegible]	Wartenberg	0.917	0.969	1.611	36.509	0.053	0.280	0.174	0.452	0.046	1.693	3.750
[illegible]	Brienne	0.867	0.833	0.903	2.263	0.162	-	-	-	-	-	1.667
[illegible]	Craonne	0.836	0.922	1.418	0.751	0.120	-	-	-	-	-	0.611
[illegible]	Rheims	0.950	0.538	0.371	0.731	0.176	-	-	-	-	-	2.308
[illegible]	Toulouse	0.806	0.921	1.520	0.922	0.133	-	-	-	-	-	0.632
[illegible]	Kerygaom	0.976	0.725	0.316	62.475	0.088	0.176	0.557	0.056	0.022	1.392	25.000
[illegible]	Warsaw	0.825	0.700	0.792	2.507	0.262	-	-	-	-	-	2.000
[illegible]	Alamo	0.875	0.028	0.484	115.832	0.513	0.039	0.082	0.019	0.004	0.425	22.222
[illegible]	Maharajpur	0.944	0.833	0.598	0.216	0.103	-	-	-	-	-	0.778
[illegible]	Meeanee	0.750	0.909	1.583	127.916	0.166	-	-	-	-	-	7.143
[illegible]	Sobraon	0.841	0.680	0.738	0.196	0.258	-	-	-	-	-	0.600
[illegible]	Boomplats	0.943	0.986	1.995	1.402	0.029	-	-	-	-	-	1.050
[illegible]	Kaneyri	0.973	0.938	0.667	0.840	0.042	-	-	-	-	-	1.375
[illegible]	Calafat	0.500	0.600	1.083	2.083	0.594	-	-	-	-	-	1.333

TABLE II (Concluded)

BATTLE DATA PARAMETERS

No	Battle	a	d	μ	D/A	ε	λ	ρ	δ	A	D	x_0/y_0
52	Lexington	0.993	0.933	0.347	3.456	0.020	0.015	0.043	0.005	0.008	0.028	5.357
53	Wilson's Creek	0.794	0.932	1.672	0.393	0.128	-	-	-	-		0.375
54	Cedar Mt	0.860	0.940	1.496	3.977	0.097	0.464	0.310	0.693	0.232	0.924	1.333
55	La Puebla	0.939	0.962	1.822	1.297	0.034	-	-	-	-	-	0.625
56	Pea Ridge	0.938	0.938	1.000	1.000	1.000	0.065	0.032	0.032	0.032	0.032	1.000
57	Perryville	0.809	0.788	0.955	1.725	0.225	-	-	-	-	-	1.375
58	Secessionville	0.900	0.900	1.000	9.000	0.105	-	-	-	-	-	3.000
59	Selinger's Hill	0.917	0.900	0.917	4.842	0.096	0.574	0.627	0.527	0.261	1.264	2.400
60	Shiloh	0.732	0.777	1.082	0.495	0.280	0.140	0.129	0.151	0.199	0.099	0.650
61	Gettysburg	0.627	0.739	1.156	0.971	0.376	0.188	0.163	0.217	0.191	0.165	0.952
62	Vicksburg	0.869	0.804	0.832	1.341	0.175	0.002	0.003	0.002	0.002	0.003	1.392
63	Atlanta	0.940	0.898	0.775	2.503	0.082	0.002	0.003	0.002	0.001	0.003	2.041
64	Cold Harbor	0.935	0.975	1.581	8.372	0.042	0.997	0.631	1.576	0.344	2.884	1.831
65	Kennesaw Mt.	0.967	0.990	1.815	10.673	0.018	-	-	-	-	-	1.800
66	Nashville	0.937	0.952	1.139	3.242	0.057	0.057	0.050	0.065	0.032	0.102	1.581
67	Ocean Pond	0.800	0.960	1.176	1.991	0.183	-	-	-	-	-	1.200
68	Spotsylvania	0.811	0.820	1.022	3.383	0.204	0.020	0.020	0.021	0.011	0.037	1.800
69	Five Forks	0.964	0.974	1.162	2.933	0.031	0.016	0.013	0.018	0.009	0.027	1.474
70	Custozza	0.967	0.936	0.727	2.676	0.047	-	-	-	-	-	2.333
71	Hallue	0.959	0.975	1.278	0.517	0.033	-	-	-	-	-	0.562
72	St. Quentin	0.912	0.927	1.093	1.754	0.083	-	-	-	-	-	1.212
73	Loftcha	0.925	0.653	0.502	0.448	0.181	-	-	-	-	-	1.333
74	Bronkhorst-Spruit	0.953	0.402	0.330	0.736	0.203	-	-	-	-	-	0.579
75	Fish Creek	0.975	0.896	1.1092	2.435	0.121	-	-	-	-	-	1.429
76	Slivnitza	0.929	0.800	0.619	1.333	0.128	0.043	0.069	0.026	0.037	0.049	1.867
77	Atbara	0.959	0.722	0.408	0.101	0.116	-	-	-	-	-	0.778
78	Kiulienchéng	0.975	0.867	0.422	0.317	0.057	-	-	-	-	-	1.333
79	Telissa	0.966	0.856	0.502	0.494	0.074	0.037	0.073	0.018	0.052	0.026	1.400
80	Attu	0.816	0.000	0.548	6.312	0.616	0.028	0.051	0.015	0.011	0.070	4.583
81	Kwajalein South	0.943	0.139	0.336	1.965	0.300	0.050	0.149	0.017	0.036	0.070	4.175
82	Kwajalein North	0.963	0.018	0.269	2.301	0.271	0.090	0.336	0.024	0.059	0.137	5.642
83	Eniwetok	0.900	0.000	0.367	1.206	0.385	0.064	0.175	0.024	0.058	0.070	2.993

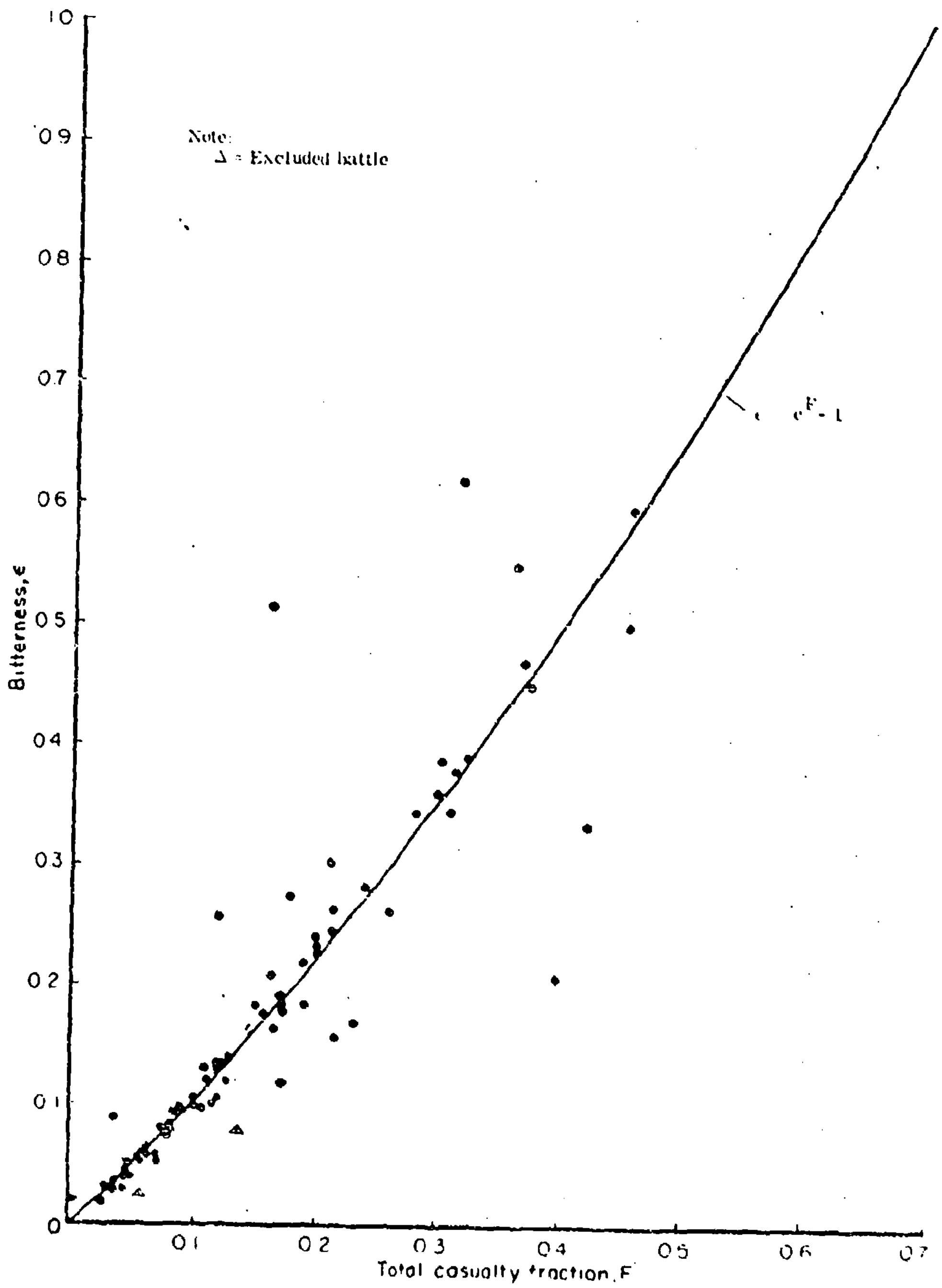


Figure 3. Linear scatter diagram of bitterness, ε, against total casualty fraction, F, for phase-2 data.

TABLE IIIa

NUMBER OF BATTLES WON BY SIDE AND BY SIGN OF ADVANTAGE, $\ln \mu$

	$\ln \mu \geq 0$	$\ln \mu < 0$	Total
Victor = D; Battle won by defending side	28	6	34
Victor = A; Battle won by attacking side	12	35	47
Total	40	41	81

Chi-square = 23.260 at 1 degree of freedom

TABLE IIIb

PERCENT OF VICTORIOUS BATTLES BY SIGN OF ADVANTAGE, $\ln \mu$, FOR VICTORIES BY EACH SIDE

	$\ln \mu \geq 0$	$\ln \mu < 0$	Total
Victor = D	82.4	17.6	100
Victor = A	25.6	74.4	100

bitterness and the like, although there are a few disturbing exceptions to the general trend. Also, as in Reference 1, we conclude that the Lanchester intensity parameter, λ, reflects the intuitive concept of combat intensity, since $\lambda = \epsilon \mid t$, where t is the duration of battle.

2. Advantage

Table IIIa shows a tabulation of the number of modified battles won by side and by the sign of $\ln \mu$, (if $\ln \mu > 0$, the defending side theoretically has the advantage; if $\ln \mu < 0$, the attacking side theoretically has the advantage). Table IIIb displays the same information expressed as a percentage of the number of victories by the respective sides.

We say that the advantage parameter, $\ln \mu$, follows the victor in a battle if the sign of $\ln \mu$ is positive when the defender wins or negative when the attacker wins. Thus, Table IIIa shows that the advantage parameter follows the victor in 63 (77.8%) of the 81 modified battles and does not follow the victor in 18 (22.2%) of the modified battles.

Treating the data of Table IIIa as a 2×2 contingency table, the value of chi square (calculated using Yates's correction) turns out to be 23.3 at one degree of freedom. This large a value of chi square would occur, if chance were the only factor affecting the data, considerably less than 0.5 percent of the time[9]. We must conclude that, beyond any reasonable doubt, some factor other than, or in addition to, pure chance has given rise to the data of Table IIIa.

A list of the advantage parameter, $\ln \mu$, values for all 83 phase-2 battles, ordered from the most extreme negative value to the most extreme positive value, was prepared and the corresponding victorious side was listed beside each advantage parameter value. This arrangement is

---

[9]See References 7 and 8 for all statistical techniques used in this paper.

[illegible] as Table IV. There is a tendency, although perhaps not quite as [illegible] as for the phase-1 data, for numerically large advantage values to follow the victor with greater fidelity than numerically small advantage values. This is further exhibited in Table Va, which gives the number of victories by side for advantage parameter values greater than 0.3, between 0.3 and -0.3, and less than -0.3. Table Vb presents the same information expressed as a percentage of the number of victories by the respective sides. From Table Va we see that, for values of ln μ numerically greater than, or equal to, 0.3, the Lanchester advantage parameter follows the victor in 39 (84%) of 46 battles, and does not follow the victor in 7 (15.2%) of these same battles.

Treating Table Va as a 2× 3 contingency table, we compute a chi-square value of 21.8 at two degrees of freedom. This large a value of chi square would occur, by chance alone, less than 0.5 percent of the time.

In sum, the Lanchester advantage parameter seems to adequately reflect the more usual, intuitive concept of probable victory.

Behavior of Individual Parameters

1. Force Ratio

Table VI shows the theoretical frequency of modified battles for various ranges of logarithmic force ratio, ln (x_0/y_0), computed on the basis of a normal distribution with the same mean and variance as the modified observed values. The observed frequency of modified battles is shown for comparison. Figure 3 exhibits the theoretical cumulative normal distributions[10] and the modified observed cumulative distribution of logarithmic force ratio. Figure 4 displays the theoretical normal frequency distribution and modified observed frequency distribution of logarithmic force ratio.

Solid line represents phase-1 results. Dashed line represents modified battle data results.

TABLE IV

ADVANTAGE PARAMETER, $\ln \mu$, AND VICTORIOUS SIDE[a]

$\ln \mu$	Victor	$\ln \mu$	Victor	$\ln \mu$	Victor
-1.313	A	-0.185	A	0.459	D
-1.132	D	-0.184	A	0.465	D
-1.109	A	-0.127	A	0.469	D
-1.091	A	-0.102	D	0.477	A
-1.058	A	-0.087	A	0.514	D
1.002	A	-0.086	A	0.524	D
0.992	A	-0.046	A	0.536	D
-0.952	A	0.000	D	0.596	D
-0.904[b]	A[b]	0.000	A	0.600	D
-0.896	A	0.022	A	0.692	A
-0.863	A	0.034	A	0.857	D
-0.726	A	0.052	D		
-0.705	A	0.079	D		
-0.695[b]	A[b]	0.080	D		
-0.689	A	0.088	D		
-0.689	A	0.089	D		
0.609	A	0.116	A		
-0.601	A	0.120	D		
-0.580	A	0.128	A		
-0.514	A	0.130	A		
-0.480	D	0.145	D		
-0.428	A	0.150	A		
-0.423	A	0.162	D		
-0.422	A	0.183	A		
-0.405	A	0.198	D		
-0.354	A	0.245	D		
-0.319	D	0.268	D		
0.304	A	0.304	D		
-0.269	D	0.304	D		
-0.259	A	0.315	D		
-0.255	A	0.329	D		
-0.242	A	0.349	A		
-0.237	A	0.403	D		
-0.233	A	0.419	A		
0.212	A	0.439	D		
-0.209	D	0.458	D		

[a] A   Attacking Side Victorious
D   Defending Side Victorious

[b] Excluded battle

TABLE Va

NUMBER OF BATTLES WON BY SIDE AND BY MAGNITUDE OF ADVANTAGE, $\ln \mu$

	$\ln \mu > 0.3$	$0.3 \geq \ln \mu > -0.3$	$-0.3 \geq \ln \mu$	Total
Victor = D	16	15	3	34
Victor = A	4	20	23	47
Total	20	35	26	81

Chi-square = 21.773 at 2 degrees of freedom

TABLE Vb

PERCENT OF VICTORIOUS BATTLES BY MAGNITUDE OF ADVANTAGE, $\ln \mu$. FOR VICTORIOUS BATTLES BY EACH SIDE

	$\ln \mu > 0.3$	$0.3 \geq \ln \mu > -0.3$	$-0.3 \geq \ln \mu$	Total
Victor = D	45.1	44.1	8.8	100.
Victor = A	8.5	42.6	48.9	100.

TABLE VI

THEORETICAL (NORMAL DISTRIBUTION) AND OBSERVED FREQUENCY OF BATTLES WITH VARIOUS VALUES OF LOGARITHMIC FORCE RATIO, $\ln x_0/y_0$

$\ln x_0/y_0$		Theoretical Frequency (Normal Distribution)	Observed Frequency	Grouped Theoretical Frequency	Grouped Observed Frequency
From	To				
$-\infty$	-0.6	8.10	3	12.15	13
-0.6	-0.4	4.05	10		
-0.4	-0.2	5.67	4	5.67	4
-0.2	0	6.48	5	6.48	5
0	0.2	8.10	11	8.10	11
0.2	0.4	8.10	16	8.10	16
0.4	0.6	8.10	8	8.10	8
0.6	0.8	7.29	7	7.29	7
0.8	1.0	6.48	4	6.48	4
1.0	1.2	5.67	2	5.67	2
1.2	1.4	4.05	2	12.96	11
1.4	∞	8.91	9		

Chi-square = 13.26 at 7 degrees of freedom

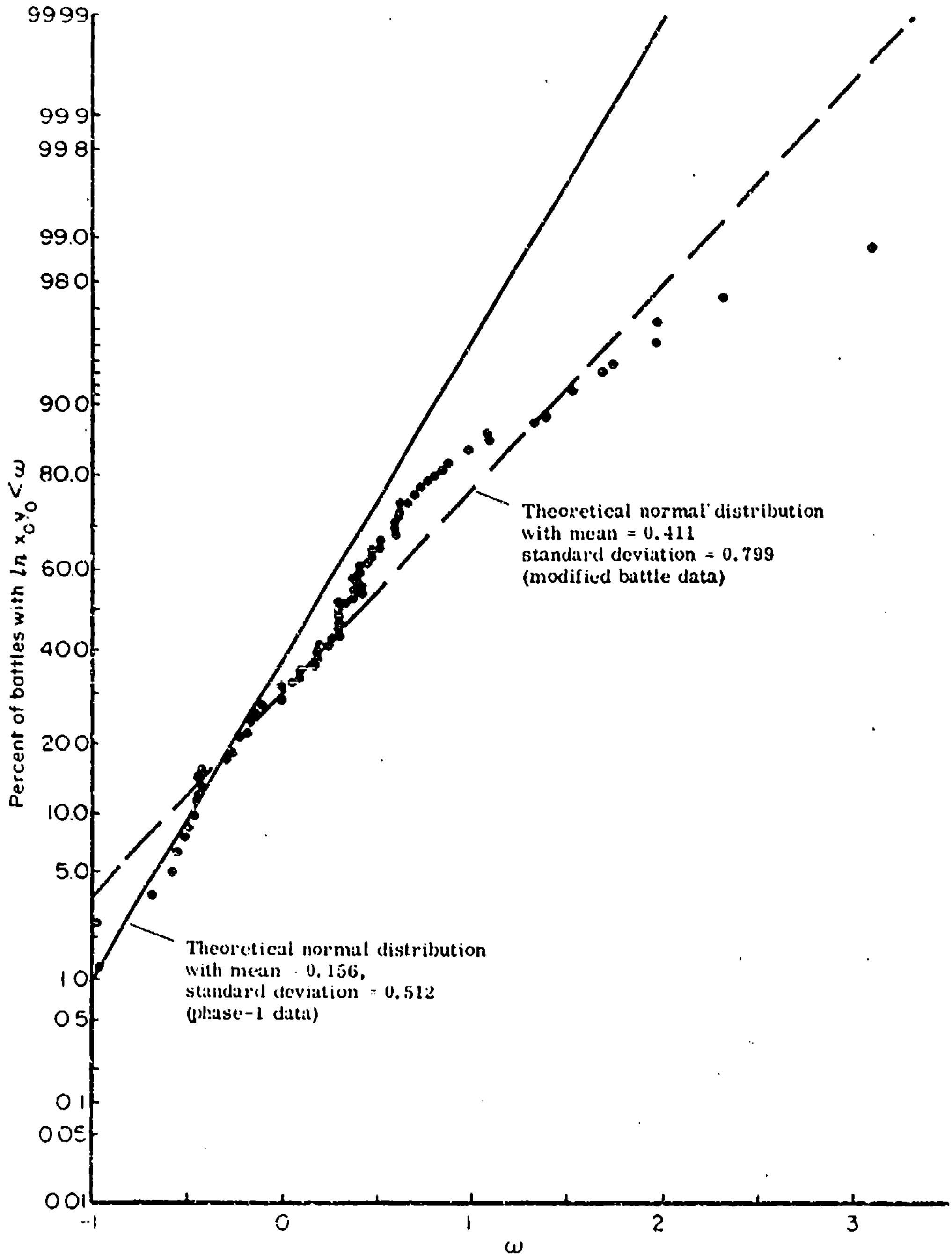


Figure 3. Theoretical and observed cumulative distribution of logarithmic force ratio, $\ln x_0/y_0$, for modified battle data.

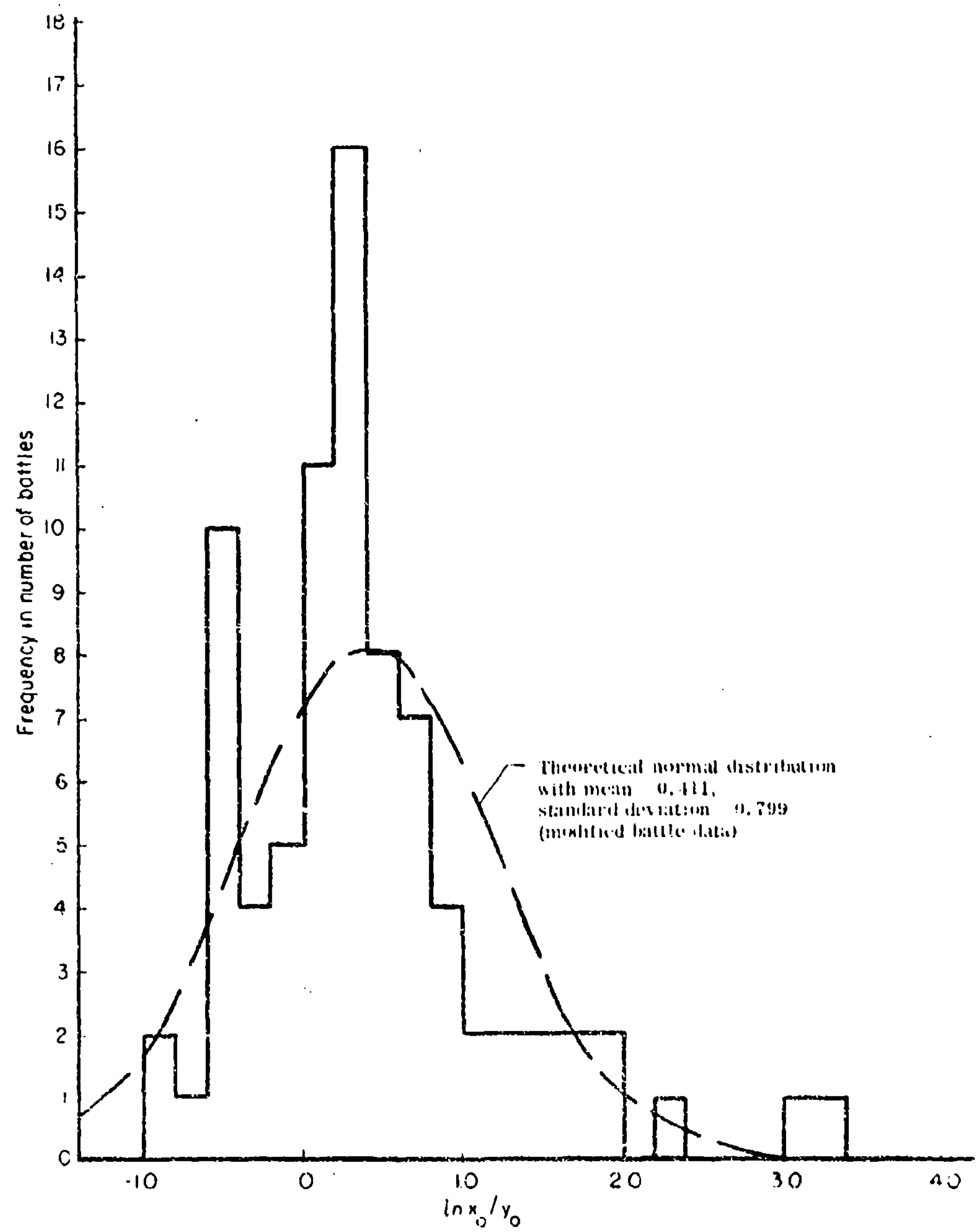


Figure 1. Theoretical and observed frequency distribution of logarithmic force ratio, $\ln x_0/y_0$, for modified battle data.

TABLE VII

THEORETICAL (NORMAL DISTRIBUTION) AND OBSERVED FREQUENCY OF BATTLES WITH VARIOUS VALUES OF DEFENDER ADVANTAGE, $\ln \mu$

$\ln \mu$ From	$\ln \mu$ To	Theoretical Frequency (Normal Distribution)	Observed Frequency	Grouped Theoretical Frequency	Grouped Observed Frequency
$-\infty$	-1.2	1.62	1		
-1.2	-1.0	1.62	5	7.29	10
-1.0	-0.8	4.05	4		
-0.8	-0.6	6.48	6	6.48	6
-0.6	-0.4	9.72	7	9.72	7
-0.4	-0.2	12.15	11	12.15	11
-0.2	0	12.15	9	12.15	9
0	0.2	12.15	16	12.15	16
0.2	0.4	8.91	7	8.91	7
0.4	0.6	5.67	12	5.67	12
0.6	0.8	4.05	2		
0.8	1.0	1.62	1	6.48	3
1.0	∞	0.81	0		

Chi-square 13.29 at 6 degrees of freedom

The degree of agreement between the theoretical normal distribution and the modified observed distribution has been tested using the chi-square goodness-of-fit test, the data being grouped as indicated in Table VI. The value of chi square obtained was 13.26 at seven degrees of freedom. This indicates that random sampling would produce a poorer fit to the theoretical distribution about seven percent of the time. This result indicates that the modified data is just barely consistent with the hypothesis that logarithmic force ratio is normally distributed. Estimated mean and standard deviation for the (modified data) logarithmic force ratio, $\ln (x_0/y_0)$, are 0.411 and 0.799, respectively.

2. Advantage

Table VII shows the theoretical (normal) and observed frequency of modified battles for various ranges of defender advantage, $\ln \mu$. Figure 5 exhibits the theoretical (normal) and the observed modified cumulative distribution of defender advantage[11].

The chi-square value computed from Table VII (grouping data as indicated by the last two columns) amounts to 13.29 at six degrees of freedom. A chi-square value this large would arise by <u>random</u> sampling only about four percent of the time. Hence the modified data is just barely inconsistent with the hypothesis that $\ln \mu$ is normally distributed. Estimated mean and standard deviation of $\ln \mu$ (modified data) are -0.119 and 0.503, respectively.

3. Bitterness

Table VIII shows the theoretical and observed modified frequency of battles for various ranges of logarithmic bitterness, $\ln \mu$. Figure 6 exhibits the theoretical cumulative normal distributions[12] and the observed

---

[11] Solid line represents phase-1 results. Dashed line represents modified data results.

[12] Solid line represents phase-1 results. Dashed line represents modified data results.

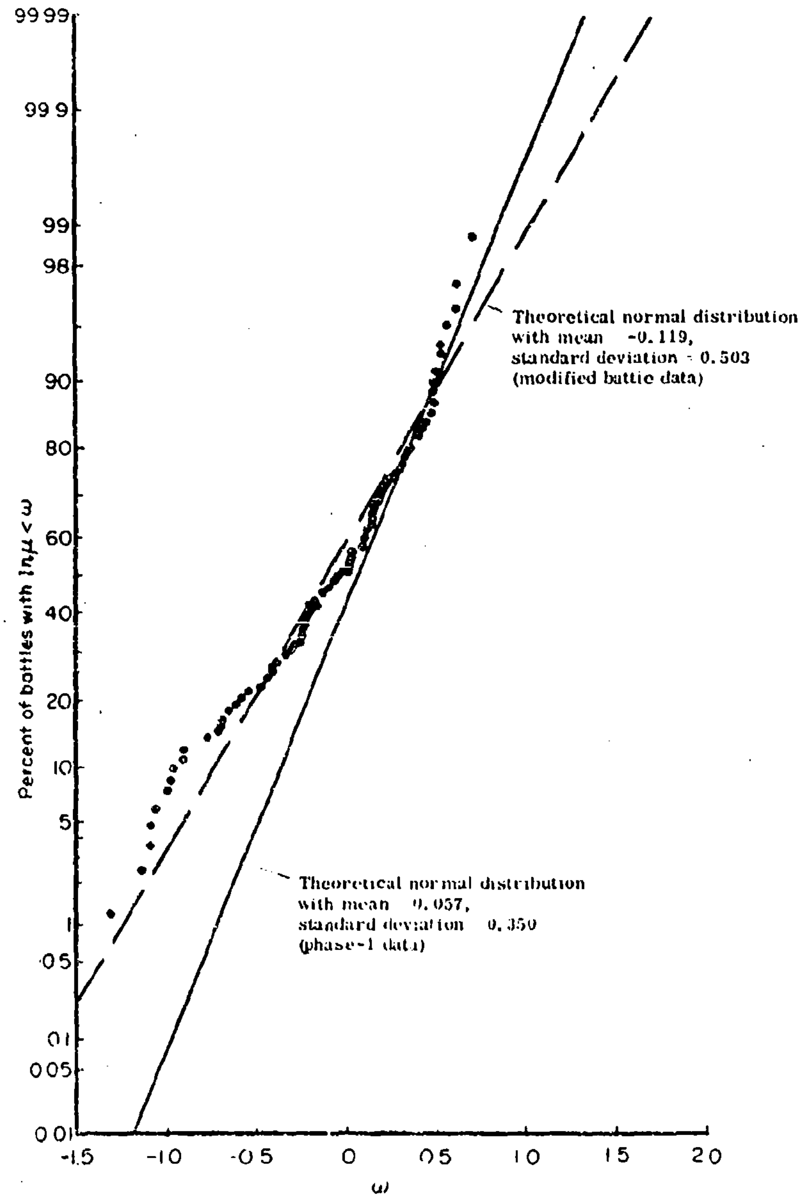


Figure 5. Theoretical and observed cumulative distribution of defender advantage, $\ln\mu$, for modified battle data.

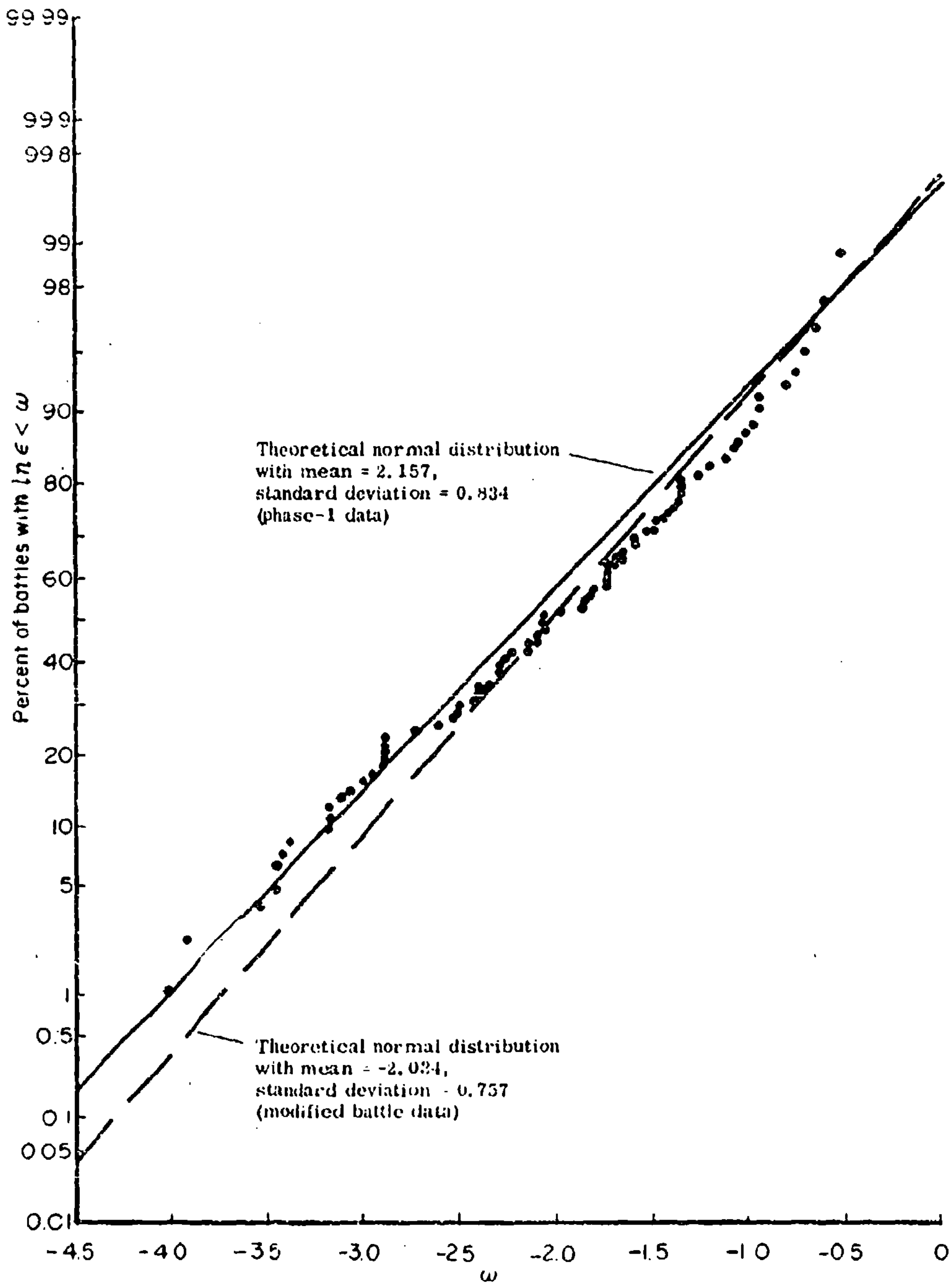


Figure 6. Theoretical and observed cumulative distribution of logarithmic bitterness, $\ln\epsilon$, for modified battle data.

TABLE VIII

THEORETICAL (NORMAL DISTRIBUTION) AND OBSERVED FREQUENCY OF BATTLES FOR VARIOUS RANGES OF LOGARITHMIC BITTERNESS, $\ln \epsilon$

$\ln \epsilon$		Theoretical Frequency (Normal Distribution)	Observed Frequency	Grouped Theoretical Frequency	Grouped Observed Frequency
From	To				
-∞	-4.0	0.40	1		
-4.0	-3.6	1.22	1	12.96	19
-3.6	-3.2	3.24	5		
-3.2	-2.8	8.10	12		
-2.8	-2.4	12.56	6	12.56	6
-2.4	-2.0	16.60	16	16.60	16
-2.0	-1.6	16.20	12	16.20	12
-1.6	-1.2	11.34	14	11.34	14
-1.2	-0.8	7.29	8		
-0.8	-0.4	2.43	6	11.34	14
-0.4	0	1.62	0		
0	∞	0			

Chi-square = 8.60 at 3 degrees of freedom

modified cumulative distribution of ln μ. The chi-square goodness-of-fit test applied to the modified data (grouped as in Table VIII) yields a value of chi square of 8.60 at three degrees of freedom. This large a value of chi square would arise by chance alone only about four percent of the time. Thus, the modified data is just barely inconsistent with the hypothesis that ln μ (modified data) are -2.034 and 0.757, respectively.

4. Surviving Fraction

The average surviving fraction of attacker (averaged over 81 modified battles) is 0.864, the corresponding value for the defender is 0.768. As indicated in Table IX, there may be a tendency for the victorious side to have a larger and less variable surviving fraction than the defeated side.

Table X is presented to provide the reader with an overview of the distribution of surviving fractions based on the modified data.

Relations Between Selected Pairs of Parameters

1. Victory and Force Ratio

Table XIa gives the number of modified battles won by side and numerical superiority or inferiority (i. e., force ratio). Table XIb gives the same information as a percentage of the bottom row marginal totals. Applying the chi-square test for independence in contingency tables (using Yates's correction), we find a chi-square value of 0.005 at one degree of freedom. This value would be exceeded by chance alone about 95 percent of the time, so the data of Table XIa is consistent with the hypothesis that victory is not dependent on numerical superiority.

In an attempt to reduce any possible masking effect arising from grouping together battles with both a small and a large degree of numerical superiority, as is done in Table XIa, values of force ratio were chosen to divide the 81 modified battles into three approximately equal groups, as indicated in Table XII. The chi-square value computed from Table XII is 0.038 at two degrees of freedom, which would be exceeded in random

TABLE IX

SOME STATISTICS OF SURVIVING FRACTIONS

Victorious Side	No. of Battles	Average Surviving Fraction of Attacker, a	Standard Deviation of a	Average Surviving Fraction of Defender, d	Standard Deviation of d
A	47	0.881	0.085	0.682	0.294
D	34	0.840	0.119	0.887	0.095

TABLE X

OBSERVED DISTRIBUTION OF SURVIVING FRACTIONS

Range of Surviving Fraction Values	Number of Battles in Stated Range					
	Attacking Side Victorious (49 Battles)			Defending Side Victorious (34 Battles)		
	Surviving Fraction of Attacker, a	Surviving Fraction of Defender, d		Surviving Fraction of Attacker, a	Surviving Fraction of Defender, d	
			90			90
0.95 - 1.00	11	4	16	5	9	21
0.90 - 0.95	14	8	23	10	13	34
0.85 - 0.90	5	5	11	5	3	12
0.80 - 0.85	10	8	19	3	2	7
0.75 - 0.80	5	3	9	5	3	12
0.70 - 0.75	0	3	3	1	3	
0.65 - 0.70	1	2	3	2	0	[illegible]
0.60 - 0.65	1	0	1	2	1	4
0.55 - 0.60	0	2	2	0	0	0
0.50 - 0.55	0	1	1	1	0	1
0.00 - 0.50	0	11	12	0	0	0

TABLE XIa

NUMBER OF BATTLES WON BY SIDE AND BY VALUE OF FORCE RATIO, x_0/y_0

	$x_0/y_0 \leq 1$	$x_0/y_0 > 1$	Total
Victor = D	10	24	34
Victor = A	13	34	47
Total	23	58	81

Chi-square = 0.005 at 1 degree of freedom

TABLE XIb

PERCENT OF BATTLES WON BY SIDE AND BY NUMERICAL SUPERIORITY

	$x_0/y_0 \leq 1$	$x_0/y_0 > 1$	
Victor = D	43.5	41.4	
Victor = A	56.5	58.6	
Total	100.0	100.0	

TABLE XII

NUMBER OF BATTLES WON BY SIDE FOR THREE LEVELS OF FORCE RATIO

	Level 1 ($0.000 \leq x_0/y_0 < 0.900$)	Level 2 ($0.900 \leq x\ /y_0 < 1.500$)	Level 3 ($1.500 \leq x_0/y_0 < 00$)	Total
Victor = D	10	11	13	34
Victor = A	13	16	18	47
Total	23	27	31	81

Chi-square 0.038 at 2 degrees of freedom

...ries about 98 percent of the time. In this case, as formerly, the modified data are consistent with the hypothesis that victory is independent of numerical superiority.

With regard to extreme force ratios, of the 81 modified battles, 21 have force ratios of 2,000 or greater. The attacker is credited with victory in 13, the defender in 8 of these 21 battles. Of the 81 modified battles, 2 have force ratios of 0.500 or less, the defender being credited with victory in both. Of the 81 modified battles, 12 have force ratios of 3,000 or greater. The attacker is credited with victory in 7, the defender in 5, of these 12 battles. Of the 81 modified battles, 7 have a force ratio of 5,000 or greater. The attacker is credited with victory in 4, the defender in 3, of these 7 battles. Also, of the 81 modified battles, 3 have a force ratio of 10,000 or greater. The attacker is credited with victory in 2, the defender in 1, of these three battles. Thus, these data lend no support to the proposition that general numerical superiority is a principal factor in tactical victory.

2. Activity Ratio and Force Ratio

Figure 7 shows a linear scatter diagram of activity ratio, D/A, against force ratio, x_0/y_0. Figure 8 presents a logarithmic scatter diagram of the same data. An approximate straight-line-fit to the data of Figure 8 is also indicated [13]. As explained in Reference 1, these regression lines are obtained from those for advantage, ln μ, on logarithmic force ratio, ln x_0/y_0.

---

[13] Solid line represents phase-1 results. Dashed line represents modified data results.

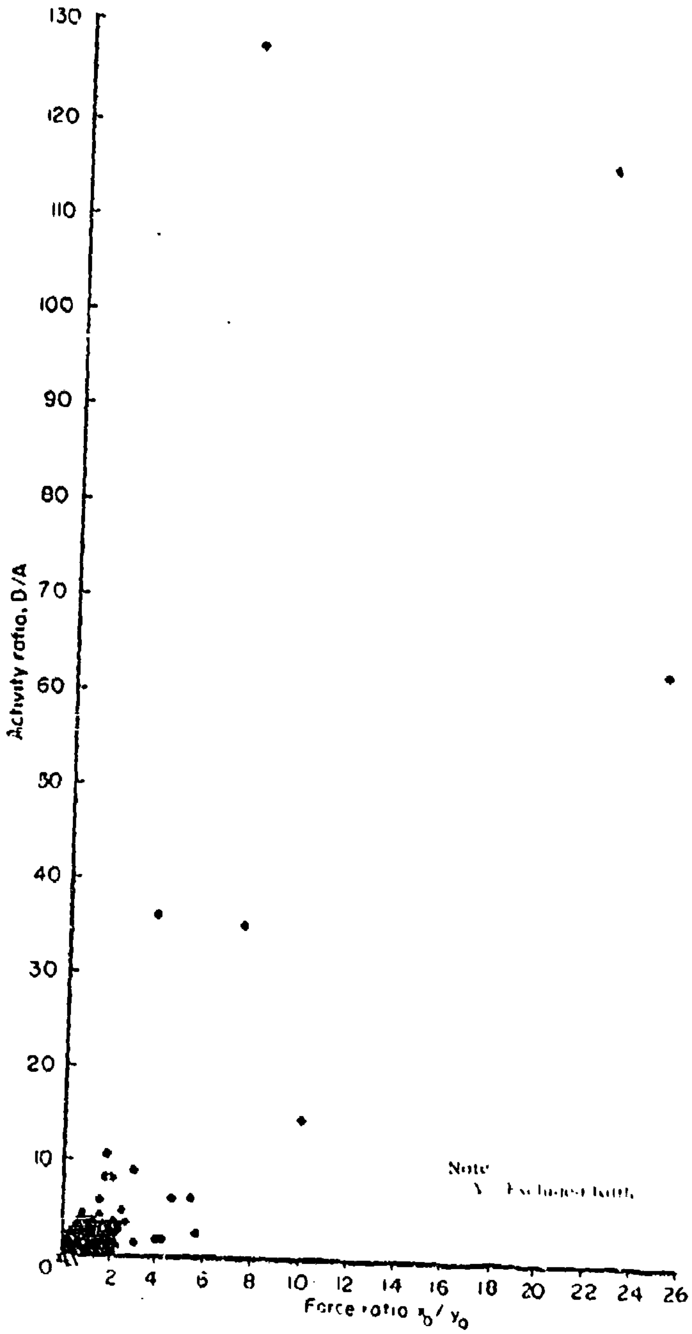


Figure 7. Linear scatter diagram of activity ratio, D/A, against force ratio, x_0/y_0, for phase-2 data.

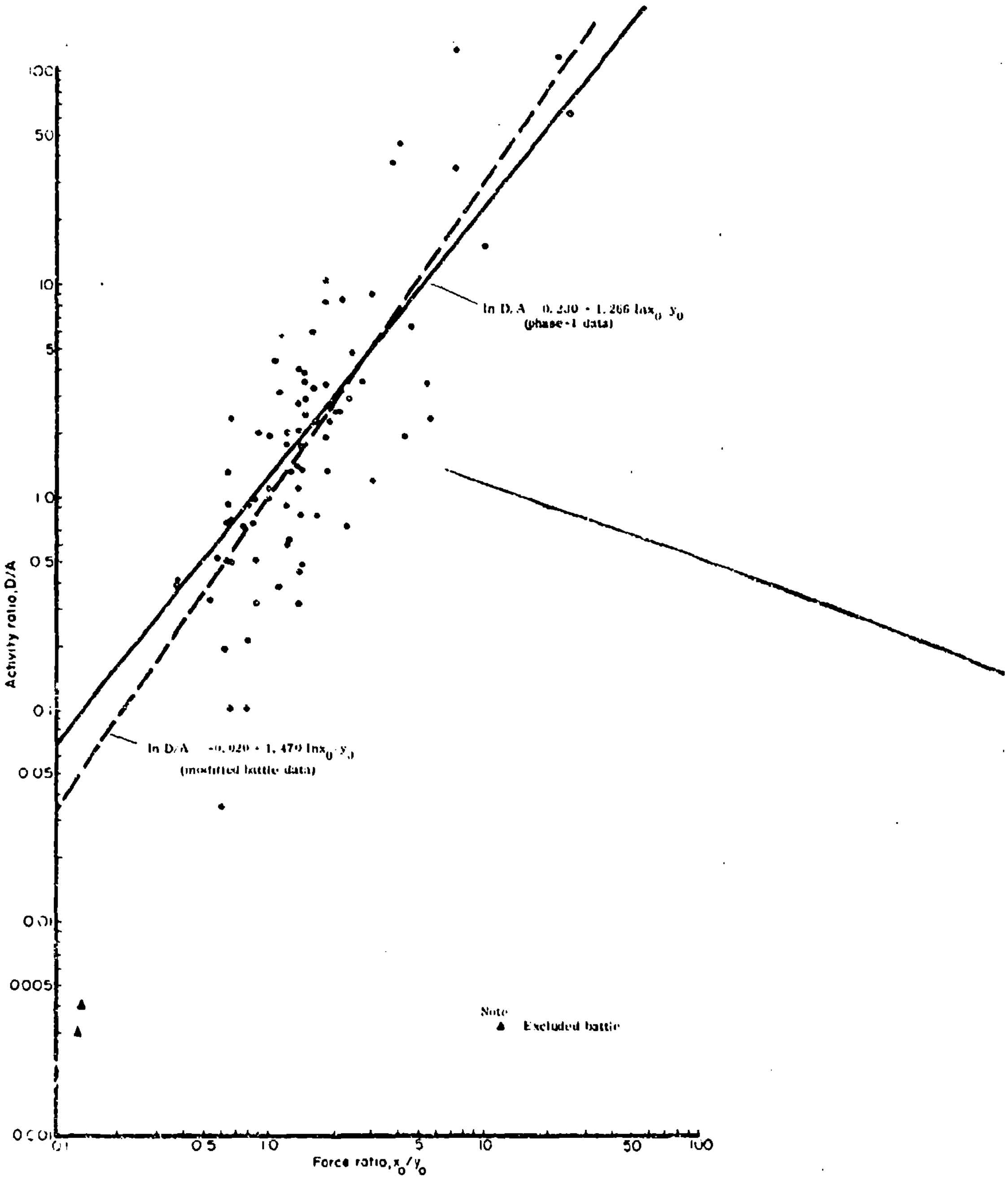


Figure 8. Logarithmic scatter diagram of activity ratio, D/A, against force ratio, x_0/y_0, for phase-2 data.

3. Defender Relative Advantage

a. Advantage and Force Ratio

Figure 9 shows a logarithmic scatter diagram at μ against force ratio, x_0/y_0, together with the regression[14] line of advantage $\ln \mu$, on logarithmic force ratio, $\ln x_0/y_0$. Table XIII displays the detailed results of the regression of (modified data) $\ln \mu$ on x_0/y_0, and these results demonstrate that (beyond reasonable doubt) defender advantage tends to decrease with increasing attacker numerical superiority[15].

As in Reference 1, the regression analysis presented here does not rigorously account for the effect of "measurement error" in the values of $\ln x_0/y_0$, nor is a more sophisticated analysis undertaken either in Reference 1, or in this paper.

To check for normality of the data about the regression line and for linearity of the regression function, we introduce the concept of residual advantage as in Reference 1:

Residual Advantage = $\ln - b - c\ln (x_0/y_0)$ where b and c are the modified data regression coefficients given in Table XIII.

Figure 10 shows the modified observed cumulative distribution of residual advantage together with two theoretical cumulative normal distributions[16]. Table XIV shows the observed and theoretical normal frequency of modified battles for various ranges of residual advantage.

---

[14] See footnote 13.

[15] See References 7 and 8.

[16] See footnote 13.

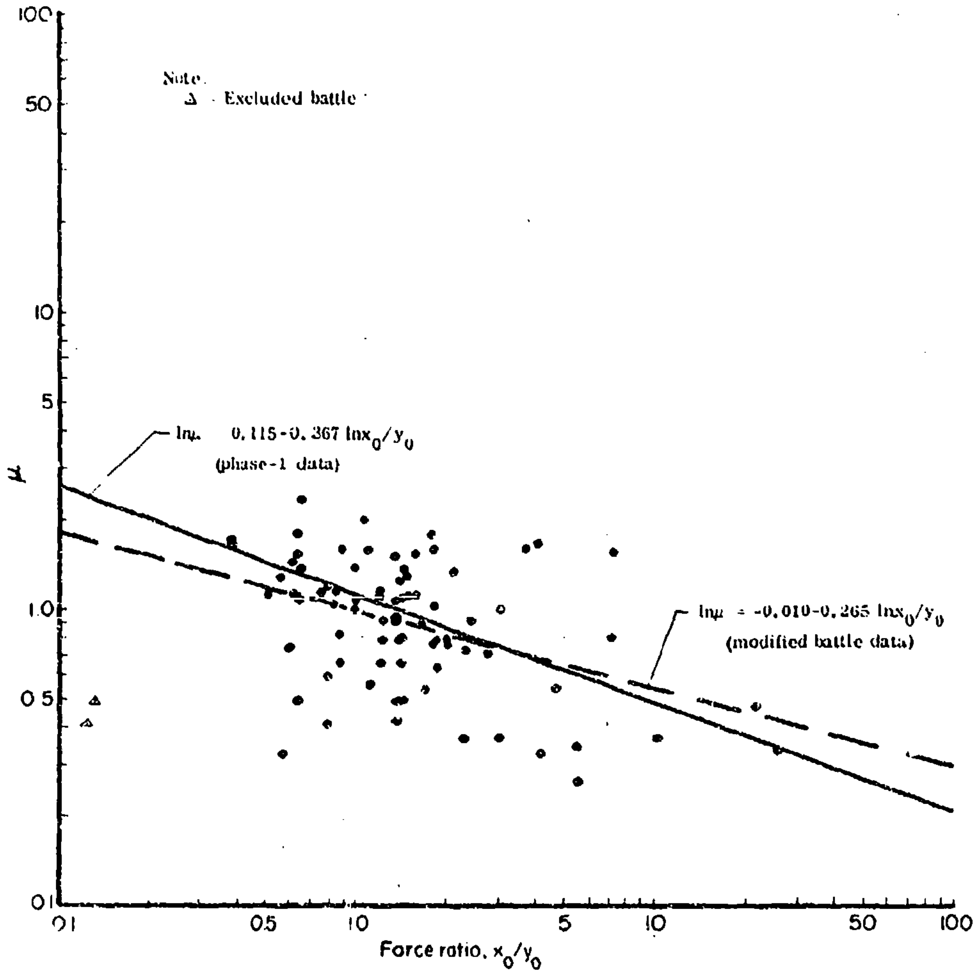


Figure 9. Logarithmic scatter diagram of μ against force ratio, x_0/y_0, for phase-2 data.

TABLE XIII

RESULTS OF REGRESSION OF ADVANTAGE, $\ln\mu$, ON LOGARITHMIC FORCE RATIO, x_0/y_0

Regression line: $\ln \mu = b + c \ln x_0/y_0$

Number of data points: 81

Estimates value of regression coefficients ±95% confidence limits:

$b = -0.010 \pm 0.115$

$c = -0.265 \pm 0.129$

Standard error of estimate: $s_{\ln \mu \mid \ln (x_0/y_0)} = 0.460$

Correlation coefficient: $r = -.420$

Sample mean:

of $\ln \mu = -0.119$

of $\ln x_0/y_0 = 0.411$

Sample variance:

of $\ln \mu = 0.503$

of $\ln x_0/y_0 = 0.799$

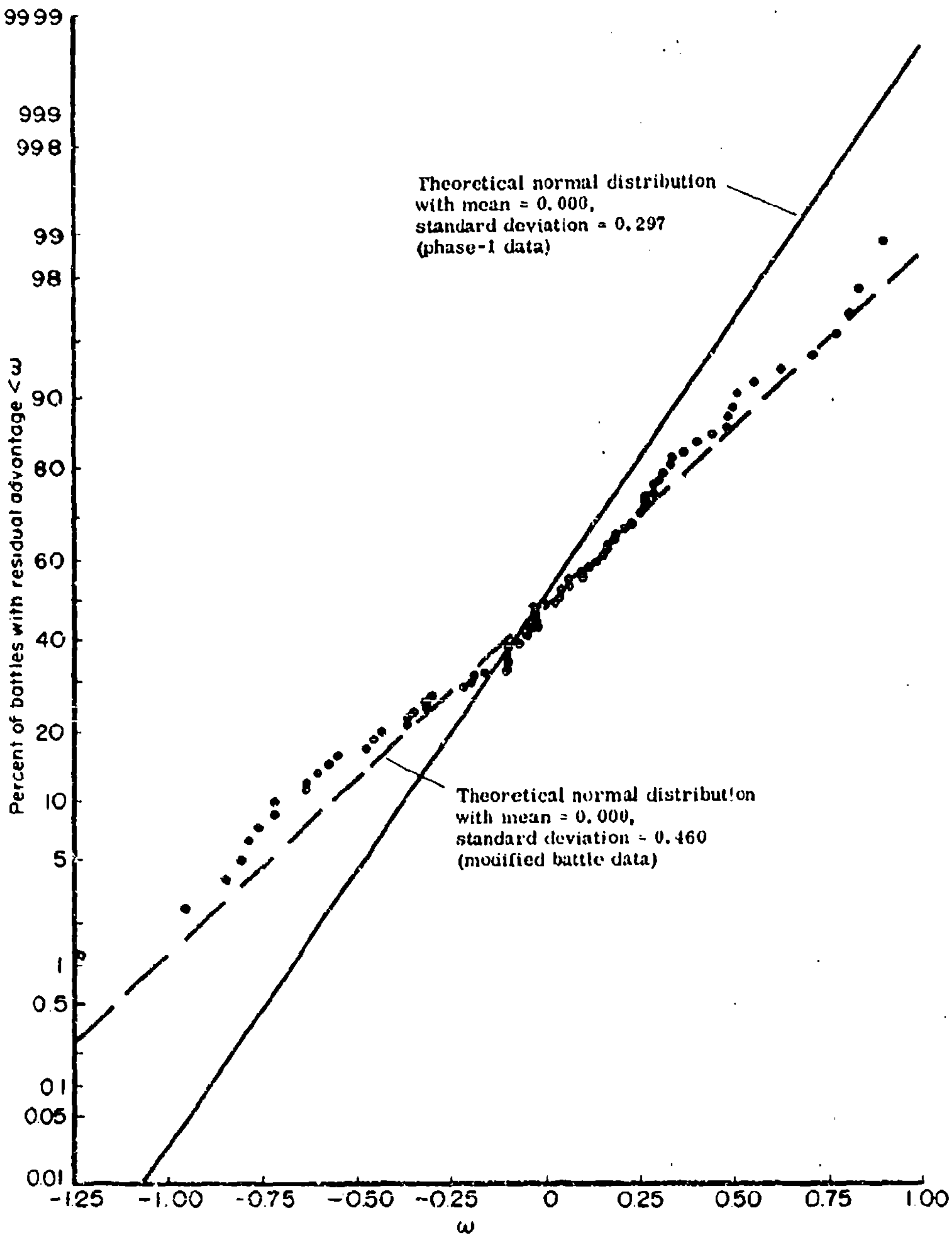


Figure 10. Theoretical and observed distribution of residual advantage for modified battle data.

TABLE XIV

THEORETICAL (NORMAL DISTRIBUTION) AND OBSERVED FREQUENCY OF BATTLES WITH VARIOUS RESIDUAL ADVANTAGES

Residual Advantage		Theoretical Frequency (Normal Distribution)	Observed Frequency	Grouped Theoretical Frequency	Grouped Observed Frequency
From	To				
$-\infty$	-0.8	3.24	4	8.10	10
-0.8	-0.6	4.86	6		
-0.6	-0.4	7.29	6	7.29	6
-0.4	-0.2	11.34	7	11.34	7
-0.2	0	13.77	16	13.77	16
0	0.2	13.77	14	13.77	14
0.2	0.4	11.34	14	11.34	14
0.4	0.6	7.29	7	7.29	7
0.6	0.8	4.86	4	8.10	7
0.8	∞	3.24	3		

Chi-square 3.45 at 5 degrees of freedom

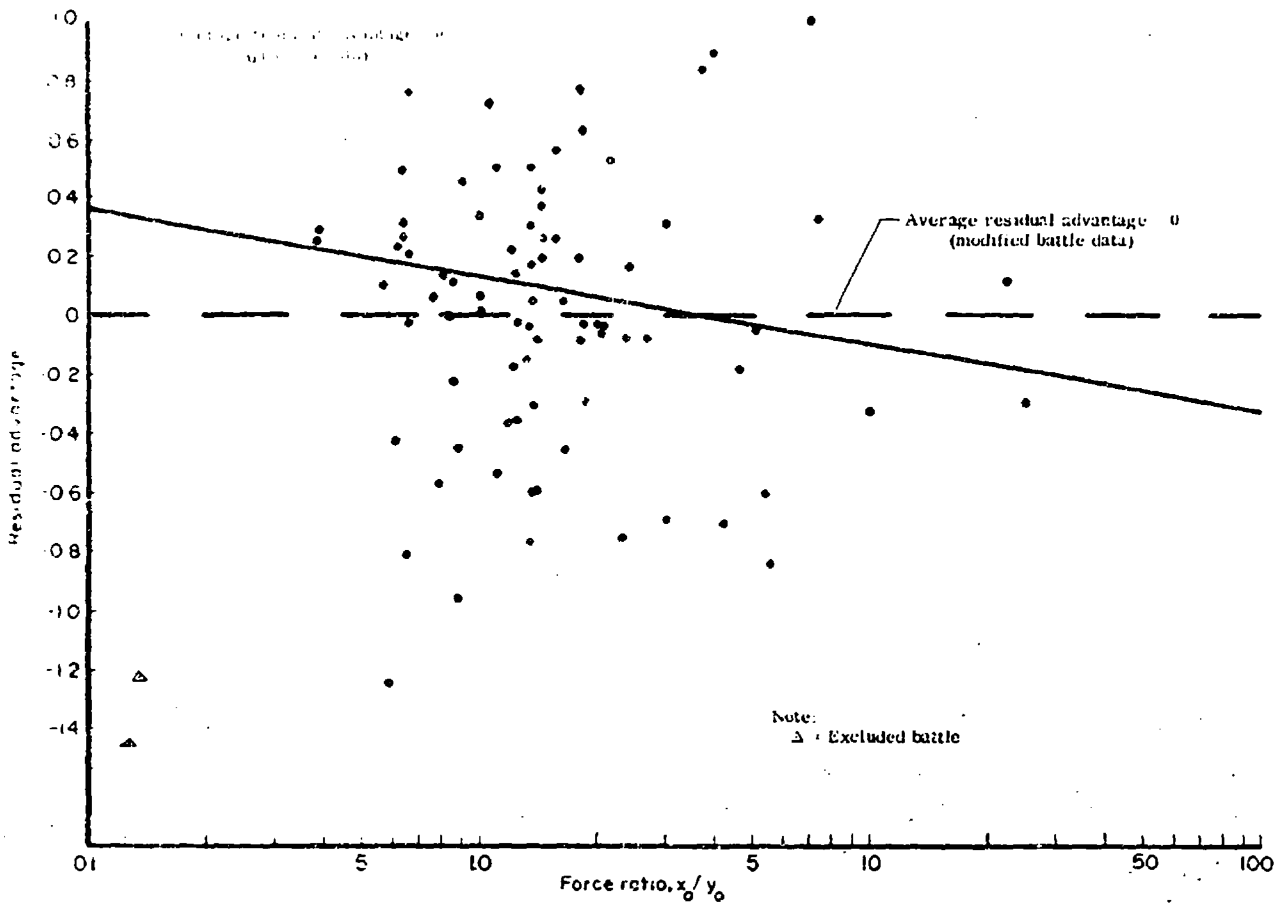


Figure 11. Semilogarithmic scatter diagram of residual advantage against force ratio, x_0/y_0, for phase-2 data.

Grouping these data as indicated in Table XIV, we compute a chi-square goodness-of-fit value of 3.48 at five degrees of freedom which indicates that a poorer fit would occur by chance alone more than 50 percent of the time. Thus, the modified data are consistent with the hypothesis that residual advantage is normally distributed.

As in Reference 1, the run test (see, e. g., Ref. 7 for description of the run test) is used as a rough check on linearity of the regression function. A count shows that, of the 81 modified battles, 42 are represented by points above the regression line of advantage on logarithmic force ratio and 39 are represented by points which lie below. A total of 42 runs occur in the modified sample of 81 residual advantage values, in order of increasing force ratio. Chance fluctuations about a linear regression line would produce fewer than 42 runs much more than 5 percent of the time, so that modified data is consistent with the hypothesis of a linear population regression curve.

The semi-logarithmic scatter diagram of residual advantage against force ratio given in Figure 11 also supports the hypothesis of linear population regression curve[17].

b. Advantage and Bitterness

Figure 12 shows a logarithmic scatter diagram of bitterness, ϵ, against μ. No trend is readily discernible to the eye. Nevertheless, a regression analysis of the modified data, the results of which are given in Table XV, indicates a slope which is significantly (at better than the one percent level) different from zero. The author does not believe that the formal analysis given in Table XV ought to be given too much credence,

---

[17] See footnote 13. The phase-1 line of zero residual advantage is calculated the difference of phase-1 and modified regression lines of advantage on logarithmic force ratio.

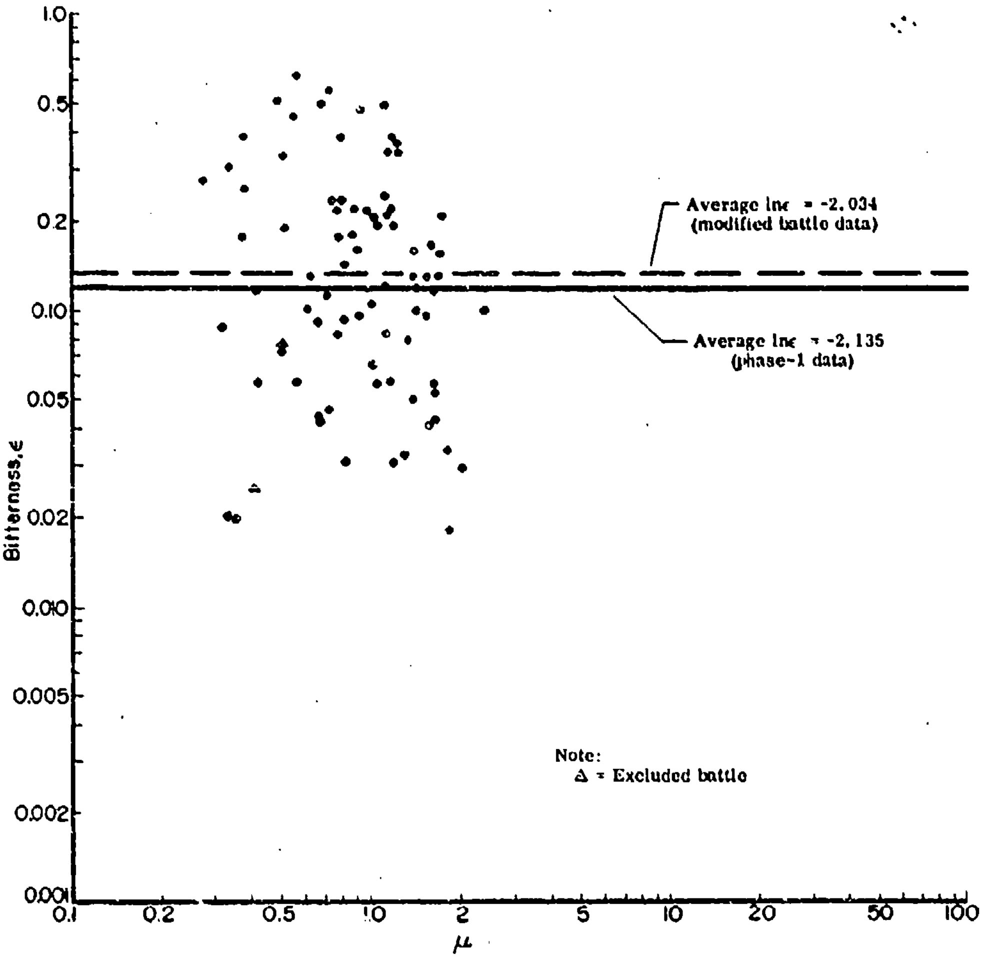


Figure 12. Logarithmic scatter diagram of bitterness, ϵ against μ for phase-2 data.

TABLE XV

RESULTS OF REGRESSION OF LOGARITHMIC BITTERNESS, $\ln \epsilon$, ON ADVANTAGE, $\ln \mu$

Regression line: $\ln \epsilon = b + c \ln \mu$

Number of data points: 81

Estimated value of regression coefficients ±95% confidence limits:

$b = -2.087 \pm 0.166$

$c = -0.448 \pm 0.323$

Standard error of estimate: $s_{\ln \epsilon \mid \ln \mu} = 0.727$

Correlation coefficient: $r = -0.298$

Sample mean:

of $\ln \epsilon = -2.034$

of $\ln \mu = -0.119$

Sample variance:

of $\ln \epsilon = 0.572$

of $\ln \mu = 0.253$

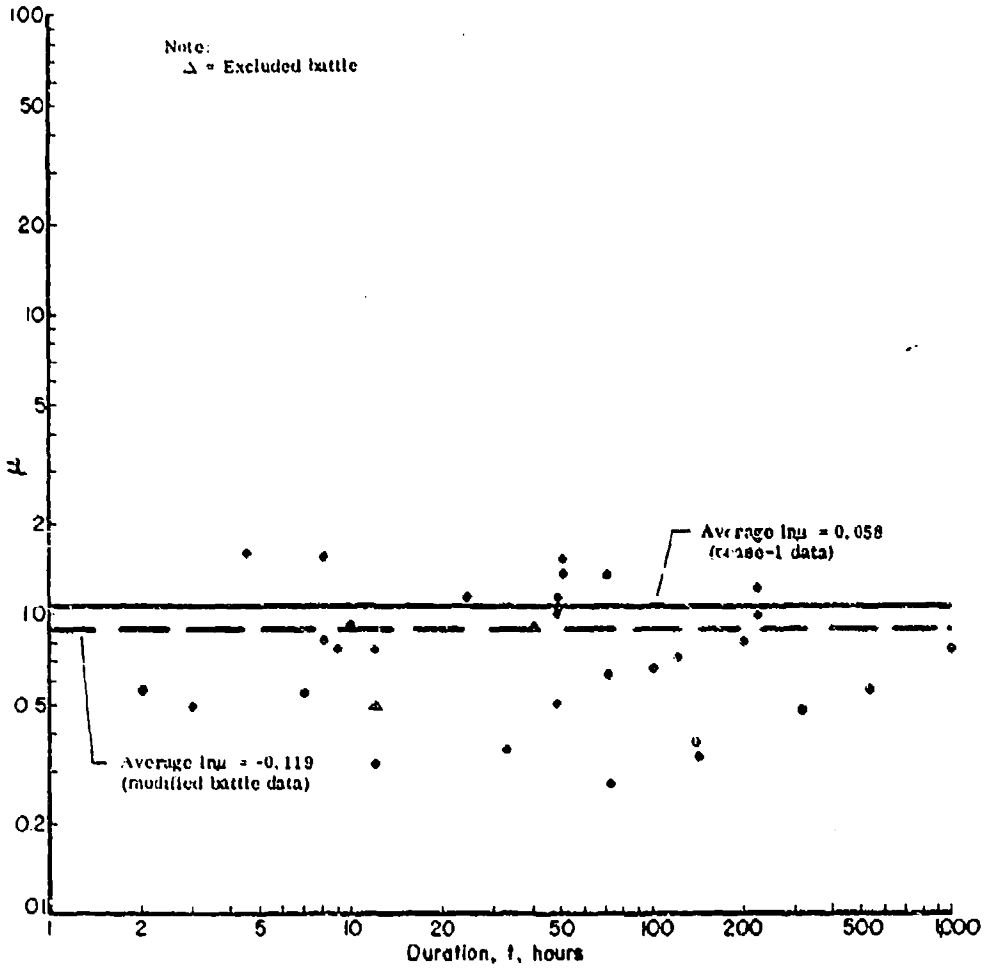


Figure 13. Logarithmic scatter diagram of μ against battle duration, t, for phase-2 data.

since so much scatter is in evidence in Figure 12 and in the wide confidence limits of Table XV. Moreover, in performing several significance tests, it often happens that some test will indicate significance as a purely chance effect (naturally, such an indication of significance is spurious and without real meaning), and the author is inclined to think that such an accidental indication of significance has occurred in the modified regression of logarithmic bitterness, $\ln \epsilon$, an advantage, $\ln \mu$. For this reason, Figure 12 shows only the lines for average logarithmic bitterness[18] rather than the computed regression line with parameters as given in Table XV.

c. Advantage and Battle Duration

Figure 13 shows a logarithmic scatter diagram of μ against battle duration, t.[19] A regression analysis of the 35 modified battles for which duration data are provided in Table I was performed, the results of which are given in Table XVI. The analytical results are consistent with the hypothesis that advantage, $\ln \mu$, is uncorrelated with logarithmic battle duration, $\ln t$.

This last result, in conjunction with the assumption of the previous section to the effect that advantage and logarithmic bitterness are uncorrelated, suggests that advantage is not correlated with logarithmic intensity, $\ln \lambda$. This suggestion is not investigated in detail.

d. Residual Advantage and Various Other Parameters

(1) Bitterness. Figure 14 gives a semilogarithmic scatter diagram of residual advantage against bitterness, ϵ. No significant degree of correlation is readily apparent to the eye.

---

[18]See footnote 13.

[19]See footnote 13.

TABLE XVI

RESULTS OF REGRESSION OF ADVANTAGE, $\ln \mu$, ON LOGARITHMIC BATTLE DURATION, $\ln t$

Regression line: $\ln \mu = b + c \ln t$

Number of data points: 35

Estimated value of regression coefficients ±95% confidence limits:

$b. = -0.006 \pm 0.364$

$c = -0.076 \pm 0.093$

Standard error of estimate: $s_{\ln \mu \mid \ln t} = 0.486$

Correlation coefficient: $r = -0.279$

Sample mean:

of $\ln \mu = -0.264$

of $\ln t = 3.350$

Sample variance:

of $\ln \mu = 0.249$

of $\ln t = 3.322$

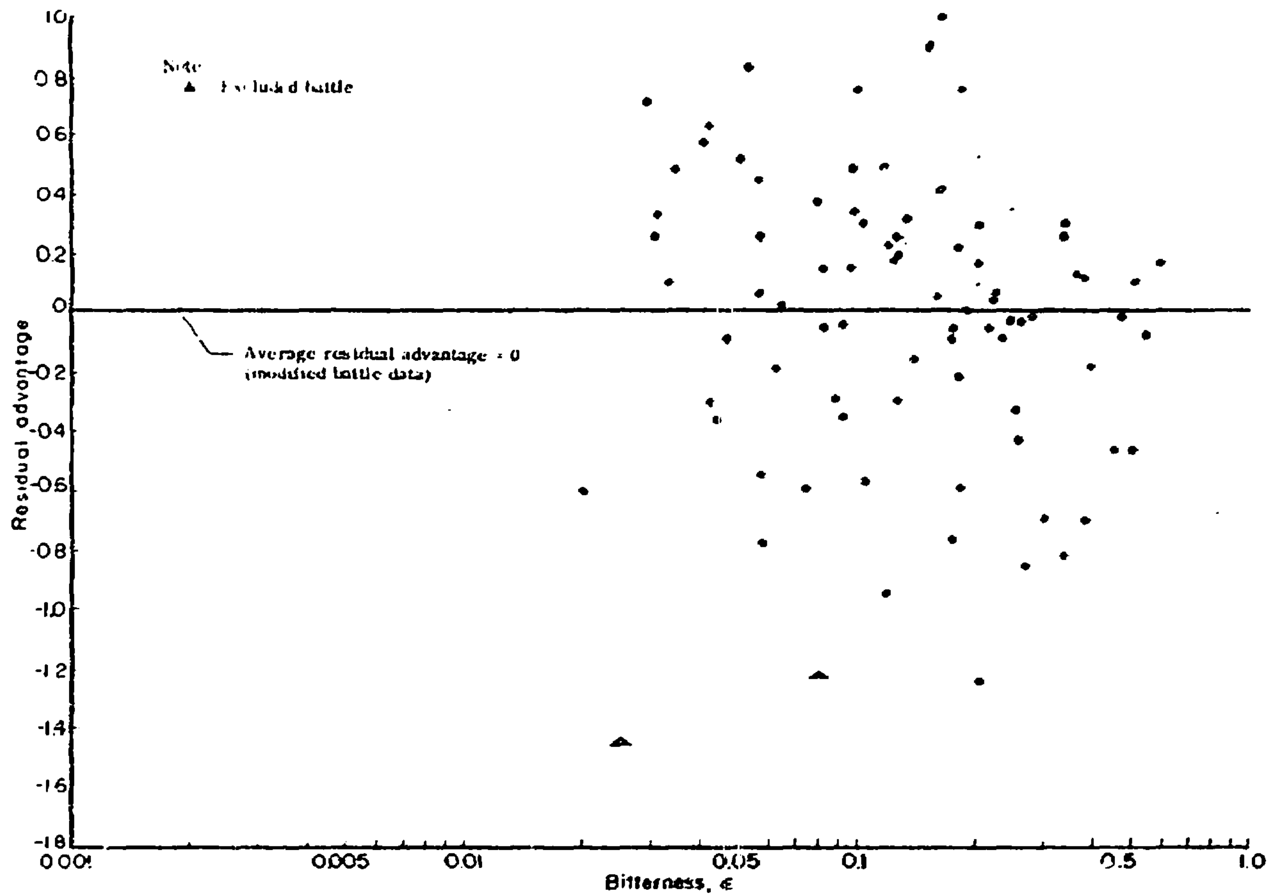


Figure 14. Semilogarithmic scatter diagram of residual advantage against bitterness, ϵ, for phase-2 data.

(2) Total Force. Figure 15 gives a semilogarithmic scatter diagram of residual advantage against total force, X. No correlation is apparent to the eye.

(3) Total Casualties. Figure 16 gives a semilogarithmic scatter diagram of residual advantage against total casualties, C . No correlation is apparent to the eye.

None of the above cases have been analyzed in detail, principally because, for the most part, the results seem to be predictable from simple inspections of the scatter diagrams.

e. Residual Advantage and Battle Date

Figure 17 shows a linear scatter diagram of residual advantage against battle date. To fit all of the battles on the same sheet, the time scale is broken at the year 1740, and the battles which occurred prior to 1740 are plotted to the left of their date of occurrence in two columns. The column on the extreme left includes all phase-2 battles which took place before 1600, the other column includes all phase-2 battles which took place between 1600 and 1740. No consiste·· trend of residual advantage with battle date is apparent in Figure 17.

f. Narratives for Residual Advantage

To provide some information on nonquantitative factors which might affect residual advantage, eight modified battles with the most extreme values of residual advantage were selected for further study ...... four battles for the most extreme positive and four for the most extreme negative values of residual advantage. In order of descending value or residual advantage, these eight battles are[20] :

[20] Numbers in parenthesis give the value of residual advantage.

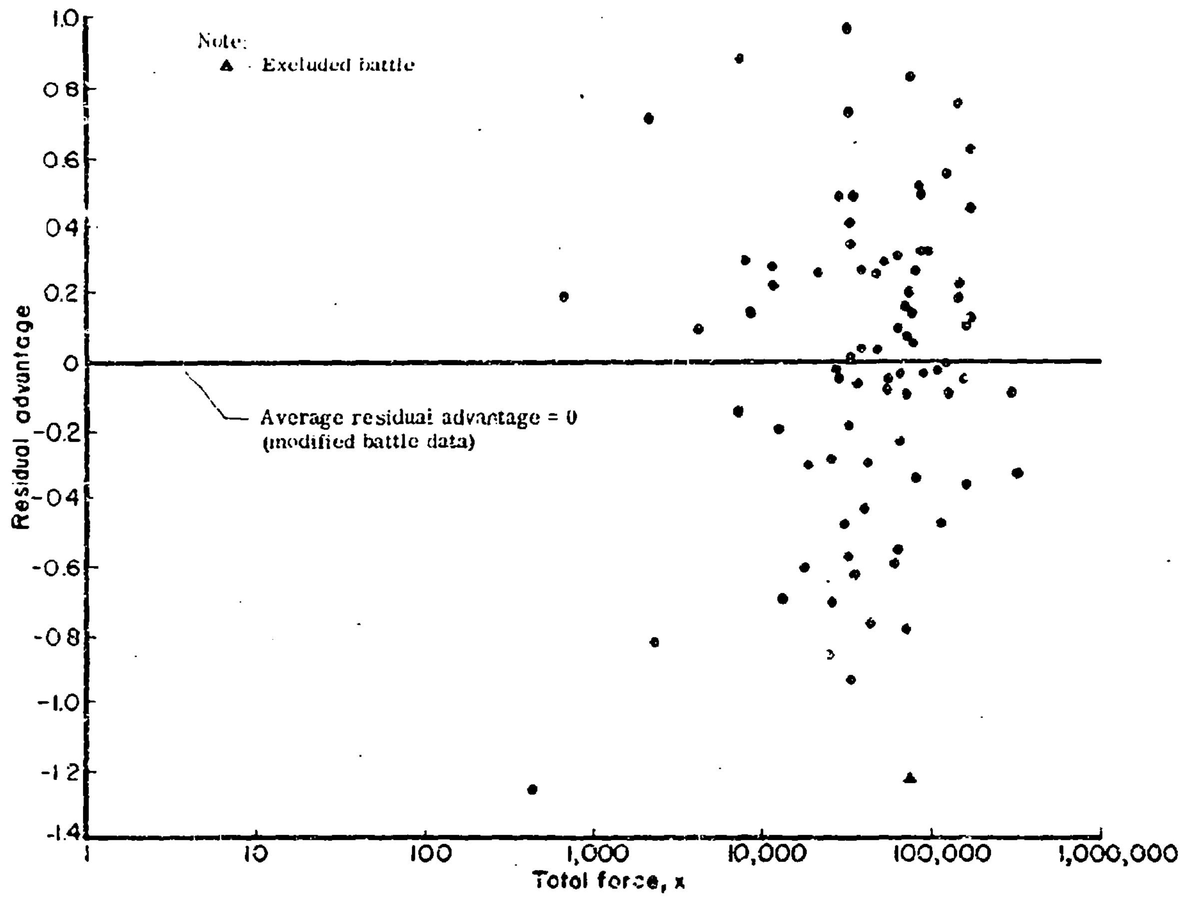


Figure 15. Semilogarithmic scatter diagram of residual advantage against total force, x, for phase-2 data.

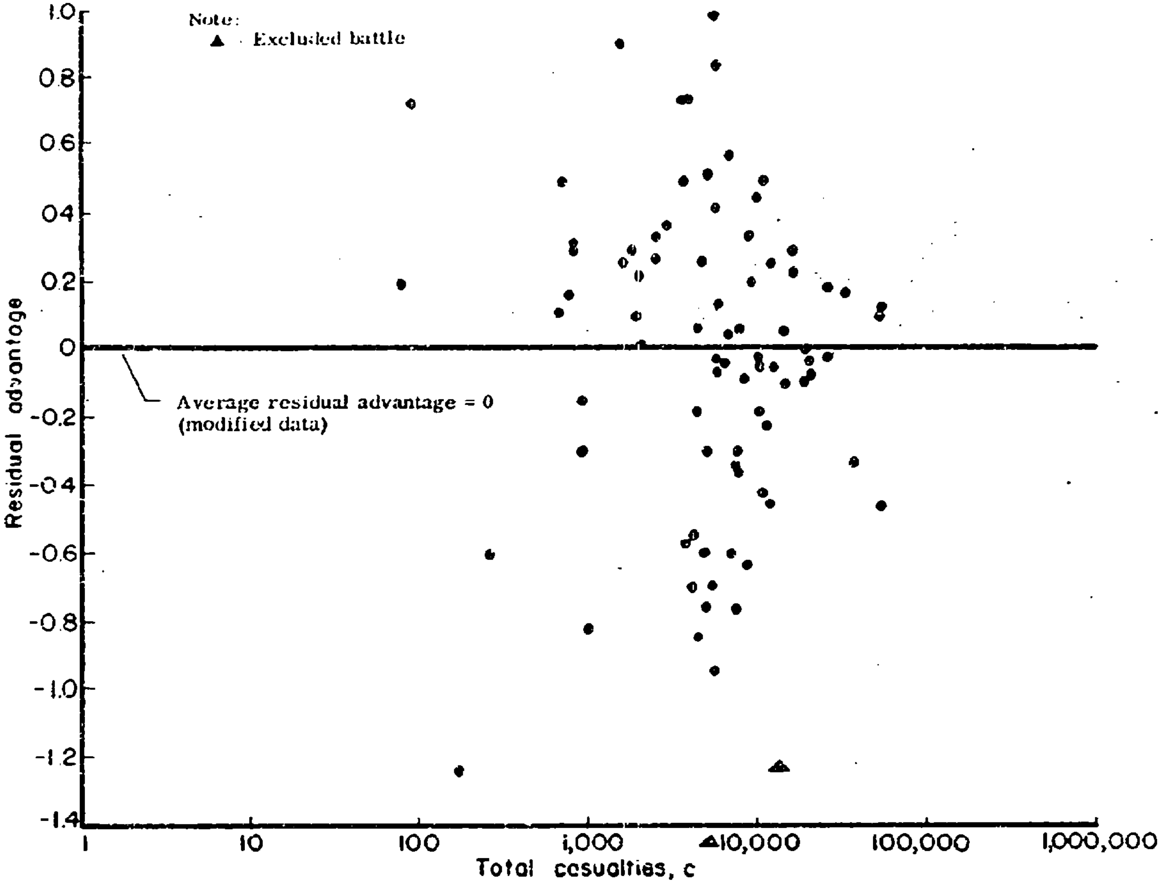


Figure 16. Semilogarithmic scatter diagram of residual advantage against total casualties, c, for phase-2 data.

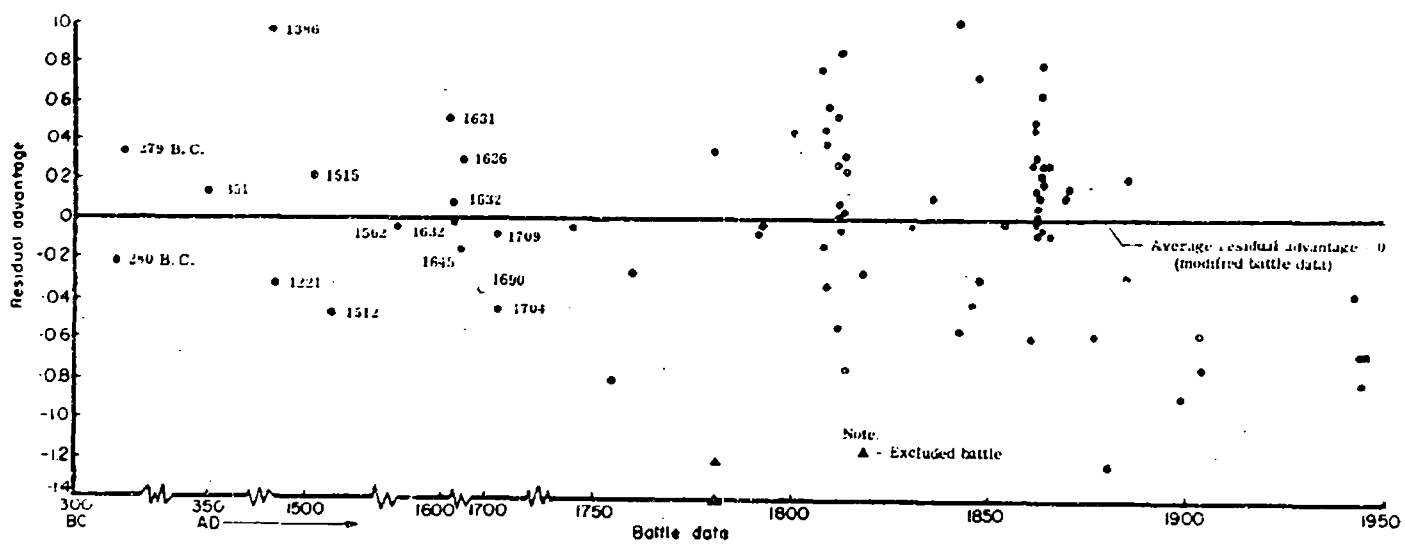


Figure 17. Linear scatter diagram of residual advantage against battle date for phase-2 data.

TABLE XVIIa

NUMBER OF BATTLES WON BY SIDE AND BY SIGN OF RESIDUAL ADVANTAGE

	Residual Advantage ≥ 0	Residual Advantage < 0	Total
Victor = D	29	5	34
Victor = A	13	34	47
Total	42	39	81

Chi-square = 23.99 at 1 degree of freedom

TABLE XVIIb

PERCENT OF VICTORIOUS BATTLES BY SIGN OF RESIDUAL ADVANTAGE FOR VICTORIES BY EACH SIDE

	Residual Advantage ≥ 0	Residual Advantage < 0	Total
Victor = D	85.3	14.7	100
Victor = A	27.7	72.3	100

Meeanee (0.9900), Sempach (0.901), Wartemberg (0.837), Kennesaw Mountain (0.762), Monongahela (-0.812), Kwajalein North (-0.845), Atbara (-0.953), and Bronkhurst Spruit (-1.244). The narrative accounts for these battles are contained in Appendix C. No pretense at completeness is made, and little real value can be gained from the sketchy description presented. There seems to be a suggestion that (1) extreme negative residual advantage values are associated with the achievement by the attacker of surprise or the delivery of heavy bombardment, and (2) extreme positive residual advantage values are associated with the choice by the attacker of direct frontal attacks lacking in surprise.

g. Residual Advantage and Victory

From Table XVIIa we see that residual advantage follows the victor in 63 (77.8%) of the 81 modified battles, and does not follow the victor in 18 (22.2%) of these battles[21]. To test the observed data against chance we use the chi-square test as before, applying Yate's correction and considering Table XVIIa as a 2 × 2 contingency table. This yields a chi-square value of 23.99 at one degree of freedom, a value which would be exceeded much less than 0.5 percent of the time if chance were the only factor affecting the data. Table XVIII lists the residual advantage values in order of increasing algebraic value, together with the corresponding victorious side. Residual advantage appears to follow the victor with at least as much fidelity as does advantage (cf., Tables IV and XVIII), especially for numerically large residual advantage values.

From Table XIXa, we see that for values of residual advantage numerically greater than 0.2, the residual advantage parameters follow the victor in 43 (84.3%) of the 51 modified battles and do not follow the victor in 8 (15.7%) of these battles. Treating Table XIX as a 2 × 3 contingency

[21] Advantage also follows the victor in 77.8 percent of the 81 modified battles; however, the set of 63 battles in which residual advantage follows the victor is not the same as the set of 63 battles in which the advantage follows the victor.

TABLE XVIII

RESIDUAL ADVANTAGE AND VICTORIOUS SIDE[a]

Residual Advantage	Victor	Residual Advantage	Victor	Residual Advantage	Victor
-1.445[b]	A[b]	-0.039	A	0.489	D
-1.244	A	-0.036	A	0.500	D
-1.224[b]	A[b]	-0.025	D	0.516	D
-0.953	A	-0.017	A	0.568	D
-0.845	A	-0.004	A	0.628	D
-0.812	A	0.010	A	0.715	A
-0.777	A	0.043	D	0.753	D
-0.760	A	0.048	A	0.782	D
-0.702	A	0.061	D	0.837	A
-0.702	A	0.062	A	0.901	D
-0.603	A	0.102	D	0.990	D
-0.603	A	0.106	A		
-0.590	A	0.113	D		
-0.571	A	0.134	A		
-0.545	A	0.150	D		
-0.464	A	0.155	A		
-0.456	A	0.166	D		
-0.429	A	0.188	A		
-0.367	A	0.193	D		
-0.361	A	0.201	D		
-0.332	A	0.220	D		
-0.311	A	0.228	A		
-0.305	D	0.255	D		
-0.289	D	0.261	A		
-0.216	A	0.263	A		
-0.188	A	0.264	D		
-0.181	A	0.286	D		
-0.151	A	0.290	D		
-0.093	A	0.301	D		
-0.086	A	0.307	A		
-0.085	D	0.327	D		
-0.084	A	0.339	D		
-0.064	D	0.373	D		
-0.056	A	0.420	D		
-0.050	A	0.448	D		
-0.041	A	0.485	D		

[a]A = Attacking Side Victorious
D = Defending Side Victorious

[b]Excluded Battle

TABLE XIXa

NUMBER OF BATTLES WON BY SIDE AND BY MAGNITUDE OF RESIDUAL ADVANTAGE

	Residual Advantage > 0.2	0.2 ≥ Residual Advantage > - 0.2	- 0.2 ≥ Residual Advantage	Total
Victor = D	22	10	2	34
Victor = A	6	20	21	47
Total	28	30	23	81

Chi-square = 26.77 at 2 degrees of freedom

TABLE XIXb

PERCENT OF VICTORIOUS BATTLES BY MAGNITUDE OF RESIDUAL ADVANTAGE FOR VICTORIES BY EACH SIDE

	Residual Advantage > 0.2	0.2 ≥ Residual Advantage > - 0.2	- 0.2 ≥ Residual Advantage	Total
Victor = D	64.7	29.4	5.9	100
Victor = A	12.8	42.5	44.7	100

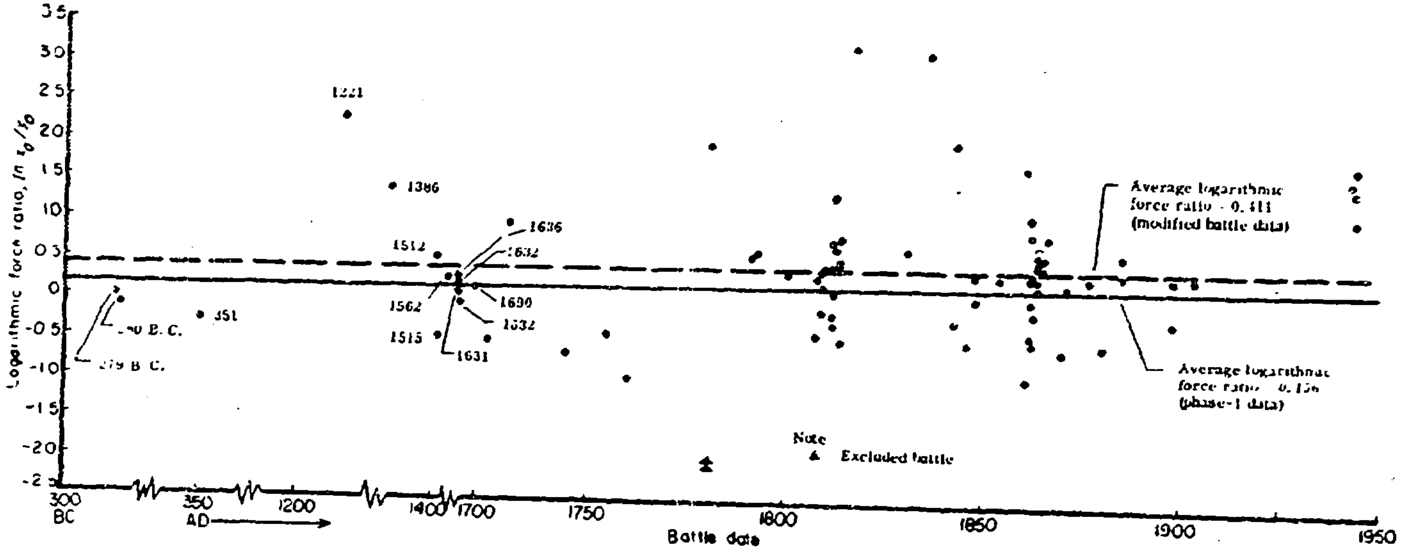


Figure 18. Linear scatter diagram of logarithmic force ratio, $\ln(x_0/y_0)$, against battle date for phase-2 data.

table, we compute a chi-square value of 26.77 at two degrees of freedom which is significantly much better than the 0.5 percent level.

As argued in Reference 1 from similar results obtained from phase-1 data, these findings indicate that the principal determinants of victory in battle are contained in the residual advantage parameter.

4. The Effects of Battle Date

The relation between residual advantage and battle date has been examined in previous paragraphs. The relation between battle date and some of the other parameters will be considered in this section.

a. Force Ratio and Battle Date

Figure 18 gives a linear scatter diagram of logarithmic force ratio against battle date. No correlation is apparent to the eye.

b. Duration of Battle Date

Figure 19 gives a semilogarithmic scatter diagram of duration in hours against battle date. There may be some tendency for battle duration to increase with battle date since about 1750.

c. Bitterness and Battle Date

Figure 20 gives a semilogarithmic scatter diagram of bitterness against battle date. To fit all of the data on one sheet, the time scale has been broken at 1550, data for battles prior to 1500 being plotted in a column at the extreme left together with the date at which the battle occurred.

There may have been a very gradual and exceedingly long-term trend toward lower bitterness values, although the sparseness of data for battles of early times makes interpretation difficult.

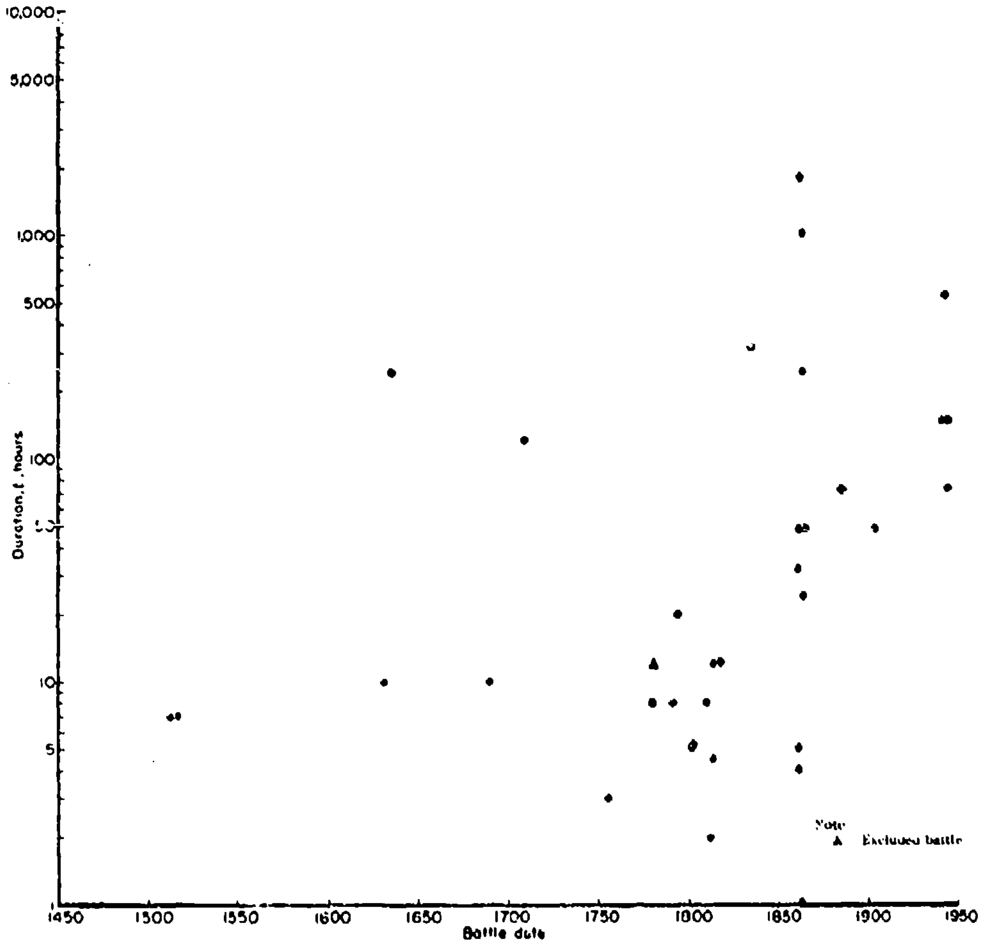


Figure 19. Semilogarithmic scatter diagram of battle duration, t, against battle date for phase-2 data.

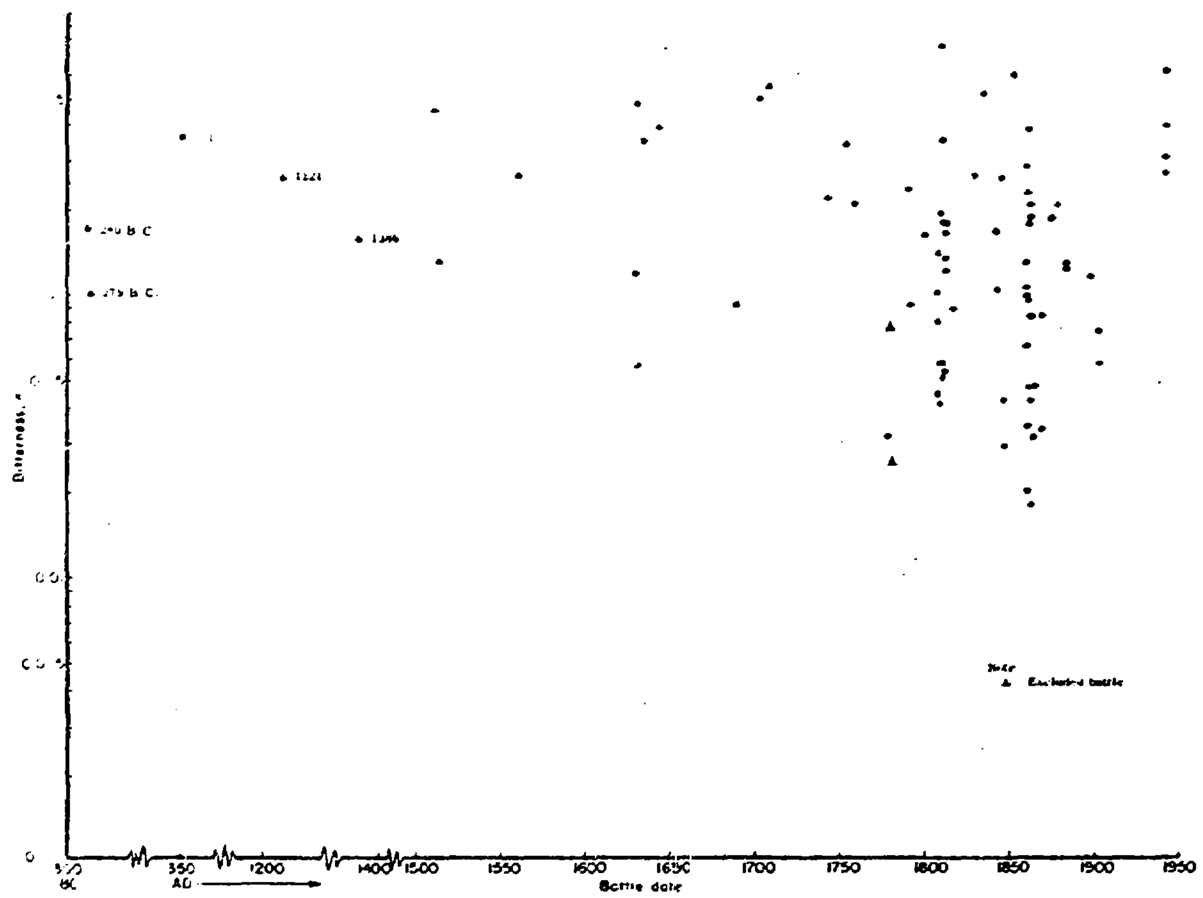


Figure 20. Semilogarithmic scatter diagram of bitterness, ϵ, against battle date for phase-2 data.

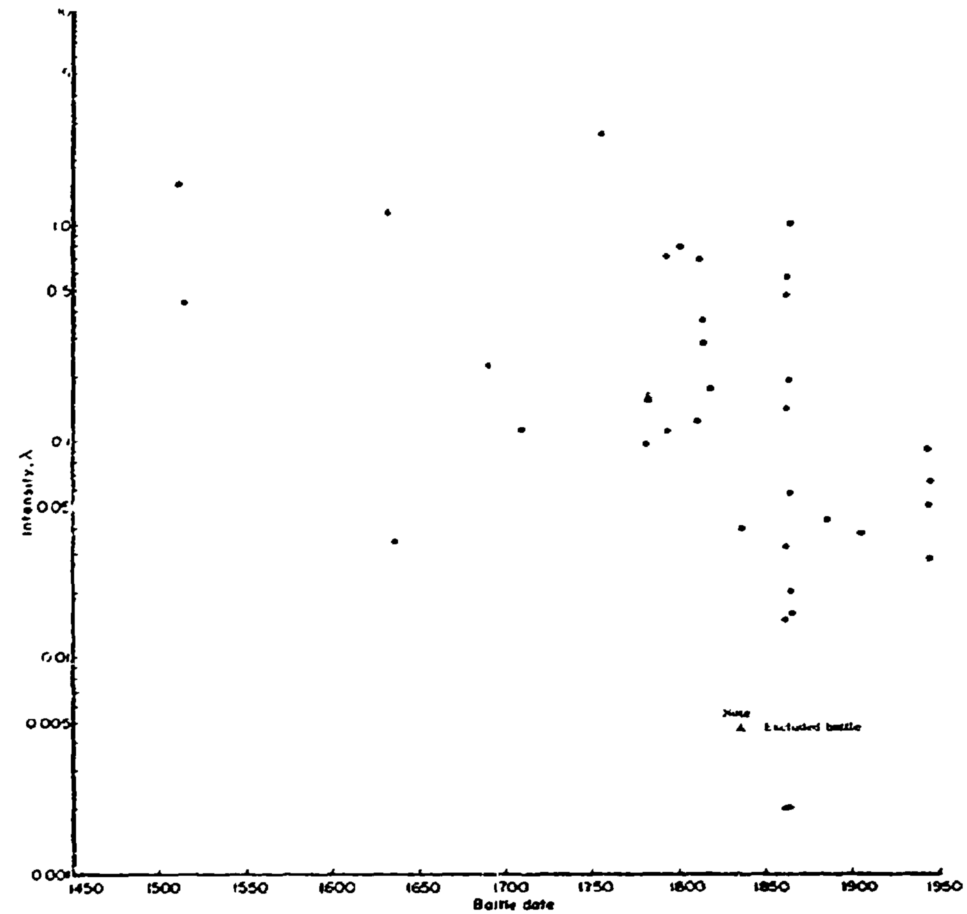


Figure 21. Semilogarithmic scatter diagram of intensity, λ , against battle date for phase-2 data.

d. Intensity and Battle Date

Figure 21 gives a semilogarithmic scatter diagram of intensity, λ, against battle date. The intensity seems to have declined from 1750 to the present.

5. Bitterness, Intensity, and Duration

Figure 22 gives a logarithmic scatter diagram of bitterness, ϵ, against duration, t, together with regression lines of logarithmic bitterness on logarithmic duration[22]. Table XX gives the results of a formal regression analysis of ln t for the modified data, and indicates that the modified regression line slope is significantly (5% level, two-tailed test) different from zero, though just barely so. The author is definitely inclined to accept this formally significant result since it appears to be supported by visual inspections of Figure 22, especially when it is recalled that phase-2 duration data have not been adjusted on the assumption of no night fighting, and since the Atlanta and Vicksburg campaigns may not be typical of operations with durations in excess of 1000 hours.

a. Residual Logarithmic Bitterness

Residual logarithmic bitterness is defined in a fashion similar to residual advantage. Table XXI shows the theoretical and modified observed frequency of battles for various ranges of residual logarithmic bitterness. Application of the chi-square goodness-of-fit test to these data, grouped as indicated in the last two columns of Table XXI, yields a chi-square value of 1.095 at three degrees of freedom, a value which would be exceeded by chance alone more than 75 percent of the time. Thus, the modified data is consistent with the hypothesis that residual logarithmic bitterness is normally distributed.

---

[22] Solid line represents phase-1 results. Dashed line represents modified data results.

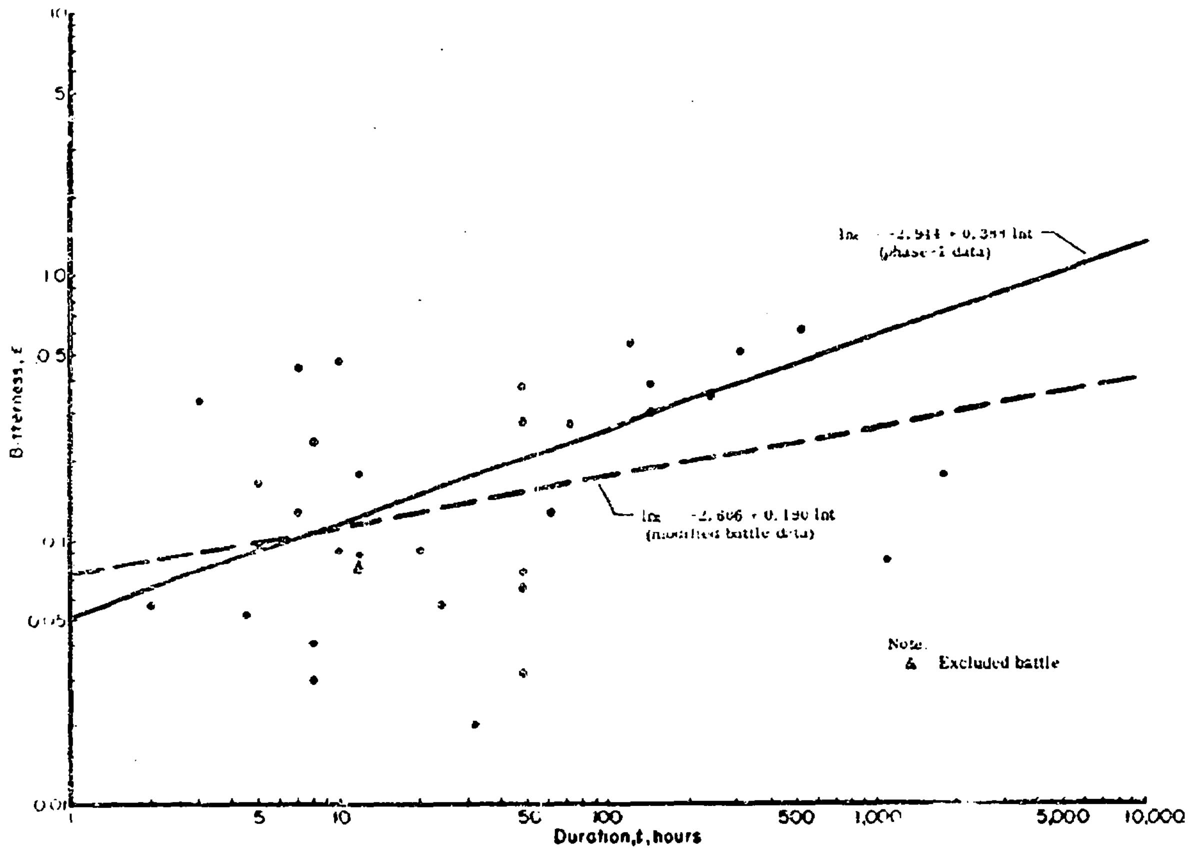


Figure 22. Logarithmic scatter diagram of bitterness, ε, against duration, t, for phase-2 data.

TABLE XX

RESULTS OF REGRESSION OF LOGARITHMIC BITTERNESS, $\ln \epsilon$, ON LOGARITHMIC DURATION, $\ln t$, for "t" in Hours

Regression line: $\ln \epsilon = b + c \ln t$

Number of data points: 35

Estimated value of regression coefficients ± 95% confidence limits:

$b = -2.606 \pm 0.643$

$c = 0.187 \pm 0.169$

Standard error of estimate: $c_{\ln \epsilon \mid \ln \mu} = 0.889$

Correlation coefficient: r = C [illegible]

Sample mean:

of $\ln \epsilon = -1.979$

of $\ln t = 3.343$

Sample variance:

of $\ln \epsilon = 0.885$

of $\ln t = 3.375$

TABLE XXI

THEORETICAL (NORMAL DISTRIBUTION) AND OBSERVED FREQUENCY OF BATTLES FOR VARIOUS RANGES OF RESIDUAL LOGARITHMIC BITTERNESS

Residual Logarithmic Bitterness		Theoretical Frequency (Normal Distribution)	Observed Frequency	Grouped Theoretical Frequency	Grouped Observed Frequency
From	To				
-∞	-1.0	4.55	4	4.55	4
-1.0	-0.5	5.60	7	5.60	7
-0.5	0	7.35	7	7.35	7
0	0.5	7.35	6	7.35	6
0.5	1.0	5.60	7	5.60	7
1.0	∞	4.55	4	4.55	4

Chi-square = 1.095 at 3 degrees of freedom

Figure 23 shows a graph of the cumulative distributions for the modified observed data and for the theoretical normal distributions[23].

Figure 24 gives a linear scatter diagram of residual logarithmic bitterness against battle date. No trend is apparent to the eye.

b. Intensity and Duration

Figure 25 gives a logarithmic scatter diagram of intensity, λ, against duration, t. Also shown are the regression lines of logarithmic intensity on logarithmic duration[24]. There is a very definite downward trend of intensity with increasing duration.

6. Total Casualties and Total Force

Figure 26 presents a logarithmic scatter diagram of total casualties, C, against total force, X. Also shown for comparison is the line $C = 0.15$, where the constant 0.15 was selected to agree approximately with the average total casualty fraction (averaged over 81 modified battles). With the exception of a few points, the line $C = 0.15X$ appears to be a reasonable fit to the data.

7. Force Ratio and Total Force

Figure 27 exhibits a logarithmic scatter diagram of force ratio, x_0/y_0, against total force, X. Also shown for comparison are the lines corresponding to the average logarithmic force ratios[25]. No consistent trend of force ratio with total force is apparent to the eye.

8. Duration, Total Force and Force Ratio

Figure 28 shows a logarithmic scatter diagram of duration, t, against total force, X, and Figure 29 shows a logarithmic scatter diagram

---

[23] See footnote 22.

[24] See footnote 22.

[25] See footnote 22.

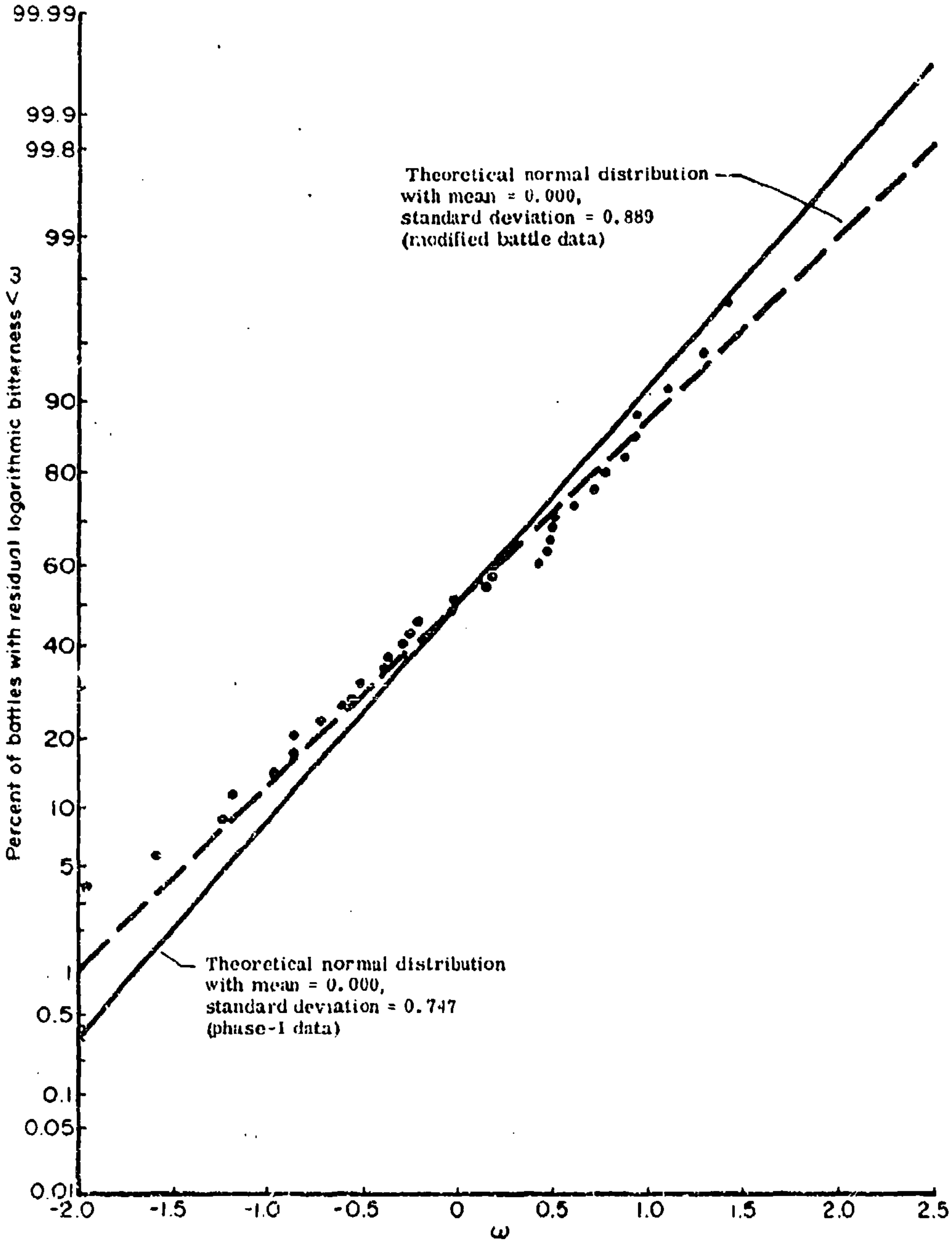


Figure 23. Theoretical and observed cumulative distribution of residual logarithmic bitterness for modified battle data.

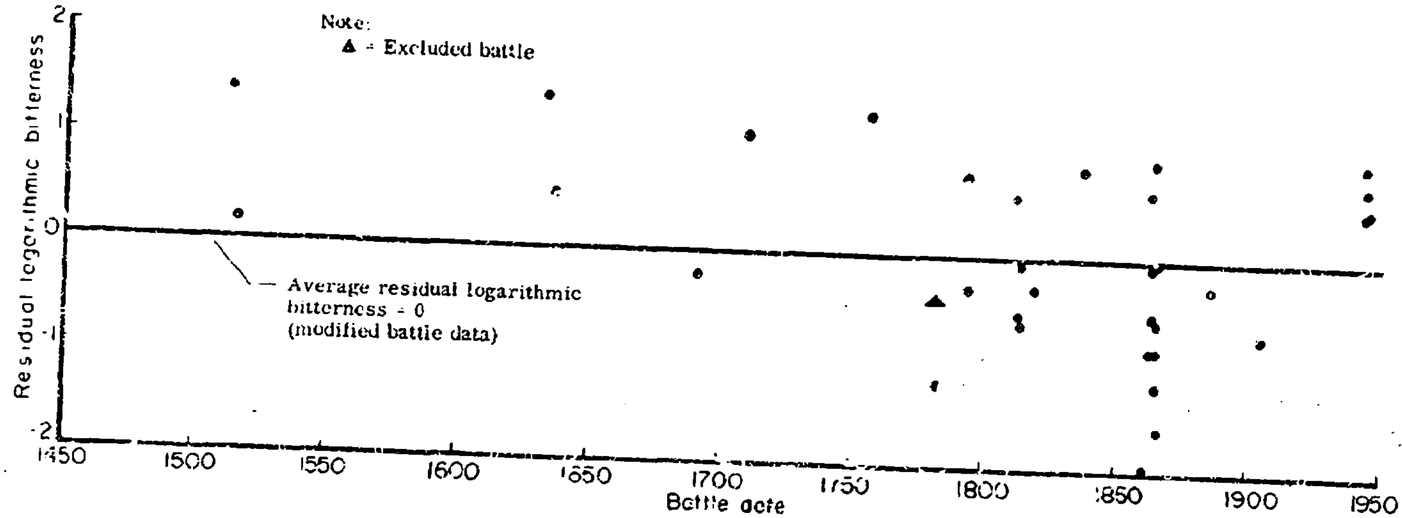


Figure 24. Linear scatter diagram of residual logarithmic bitterness against battle date for phase-2 data.

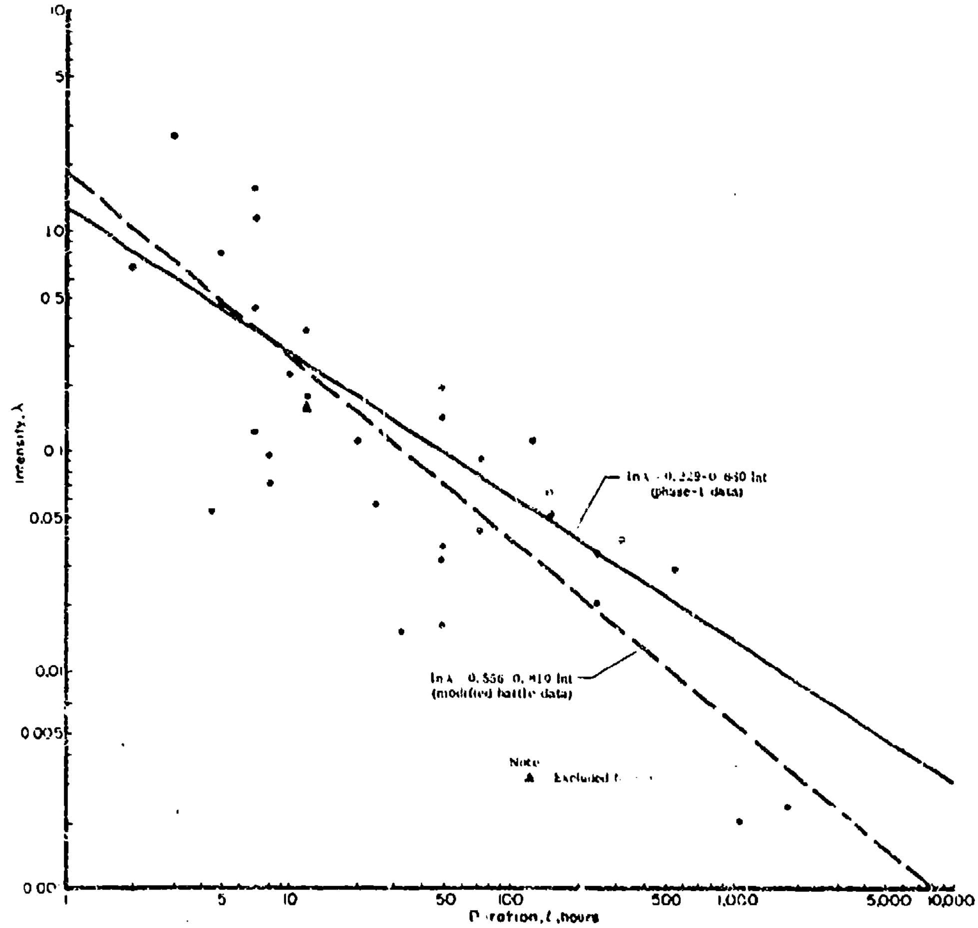


Figure 25. Logarithmic scatter diagram of intensity, λ, against duration, . . for phase-2 data.

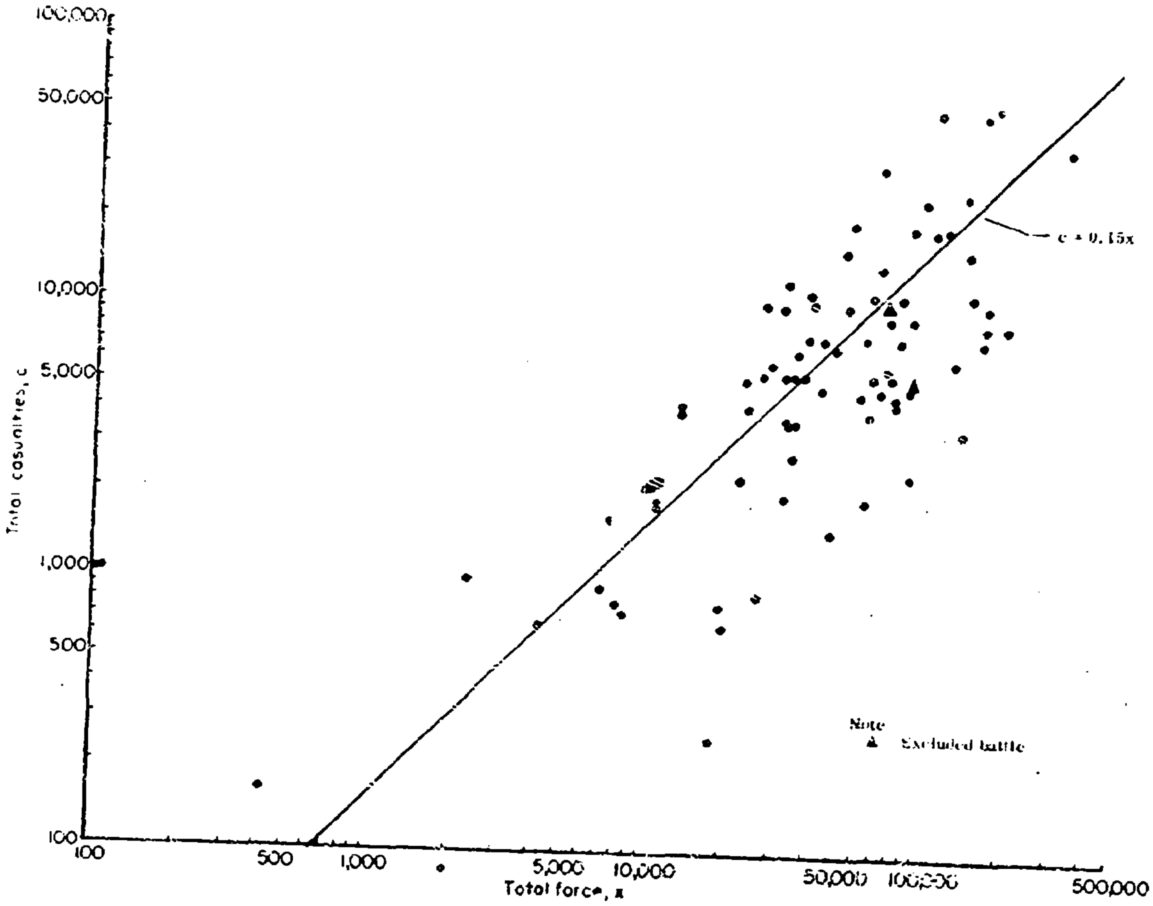


Figure 26. Logarithmic scatter diagram of total casualties, c, against total force, x, for phase-2 data.

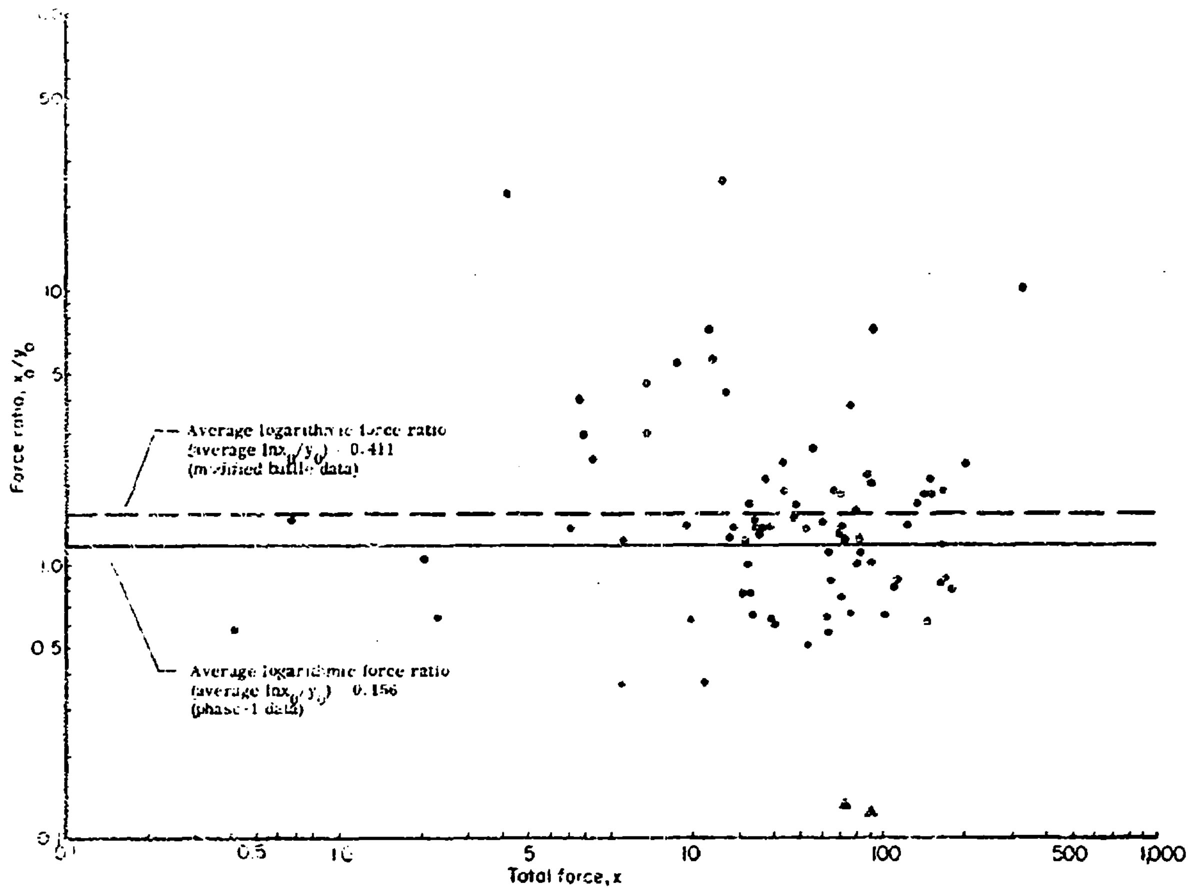


Figure 27. Logarithmic scatter diagram of force ratio, x_0/y_0, against total force, $X = x_0+y_0$, for phase-2 data.

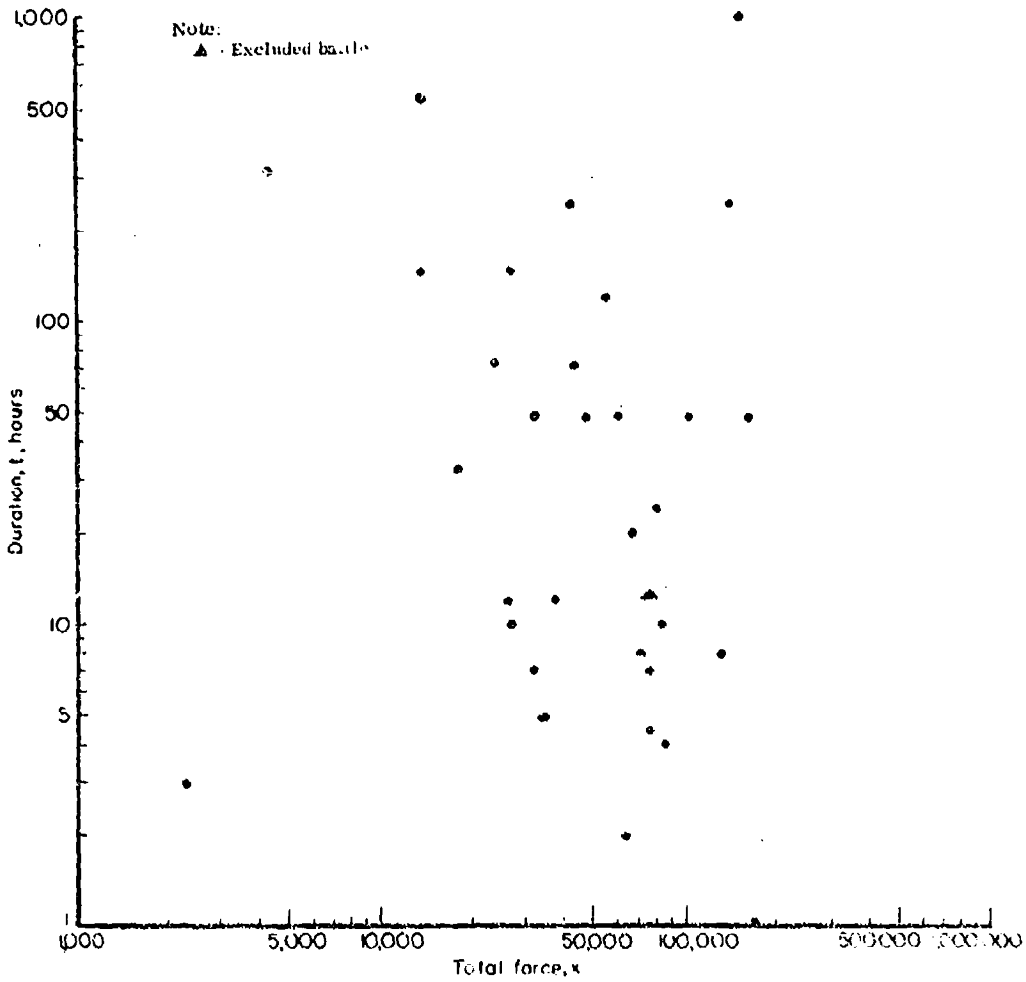


Figure 28. Logarithmic scatter diagram of duration, t, against total force, x, for phase-2 data.

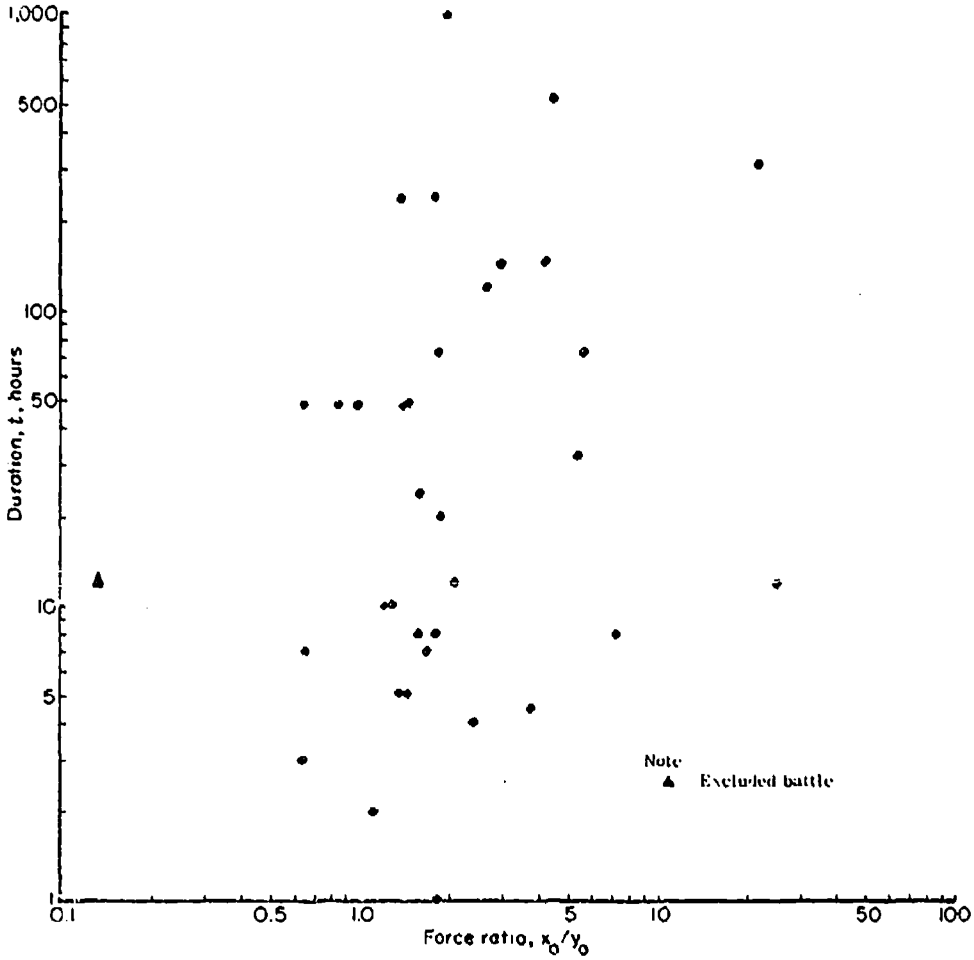


Figure 29. Logarithmic scatter diagram of duration, t, against force ratio, x_0/y_0, for phase-2 data.

of duration against force ratio, x_0 / y_0. These data have not been formally analyzed, but no consistent trends are apparent to the eye.

9. Participating Nation and Victory

The residual advantage (modified data) was arranged to represent the residual advantage favorable to each of the six most frequently participating nations[26]. Figures 30 through 35 present the results for the respective nations in the form of linear scatter diagrams of residual advantage favorable to the individual nations against battle date. Since residual advantage does not always follow the victor, a table showing victories and defeats for each nation and side was also prepared and is presented as Table XXII, which exhibits the number of victories for each participating nation as a fraction of total participation and gives approximate 95 percent confidence limits for the probability of victory.

Reference 1 notes some of the difficulties that need consideration in evaluating these data. For these reasons, no formal analysis of these data is presented here, other than that involved in estimating approximate confidence intervals for the probability of victory. Informally, it appears that no one nation has consistently been tactically superior to its opponents. For example, Table XXIII appears to indicate that the modified data are consistent with the hypothesis that the probability of victory is about 0.5. That one (i. e., about 4%) of the twenty-five 95 percent confidence intervals given in Table XXIII does not include the 0.5 probability value may be of no significance, since about five percent of such confidence intervals will exclude the true probability value by the action of chance alone.

---

[26] The residual advantage favorable to a nation participating in a battle on the defending side is the same as the residual advantage previously defined. For a nation participating on the attacking side, it is numerically equal to the residual advantage previously defined, but taken with the opposite sign.

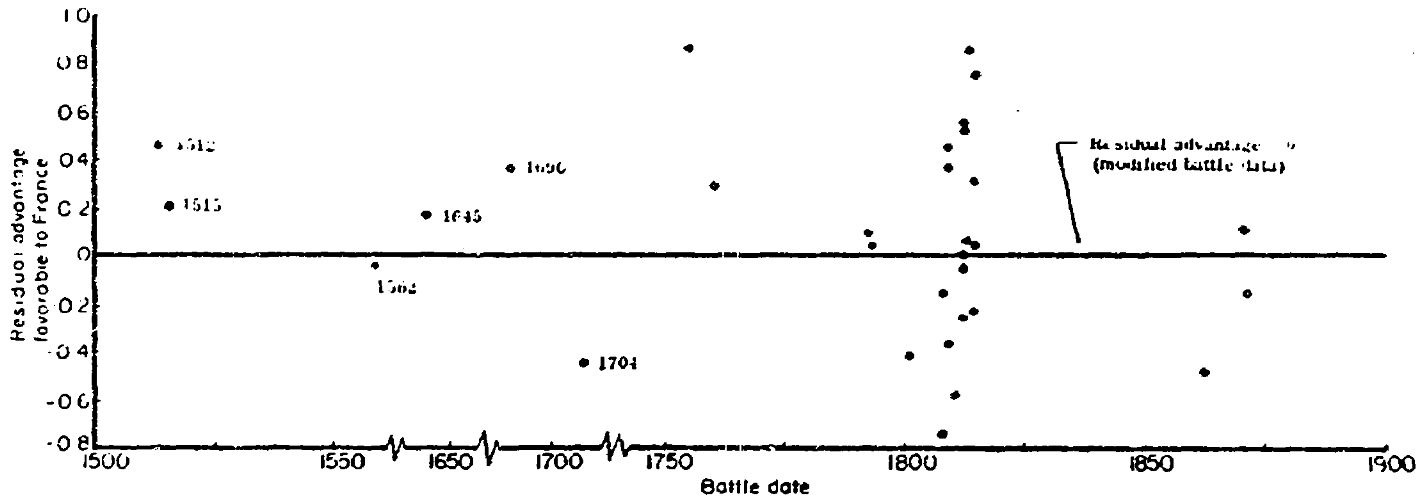


Figure 30. Linear scatter diagram of residual advantage favorable to France against battle date for phase-2 data.

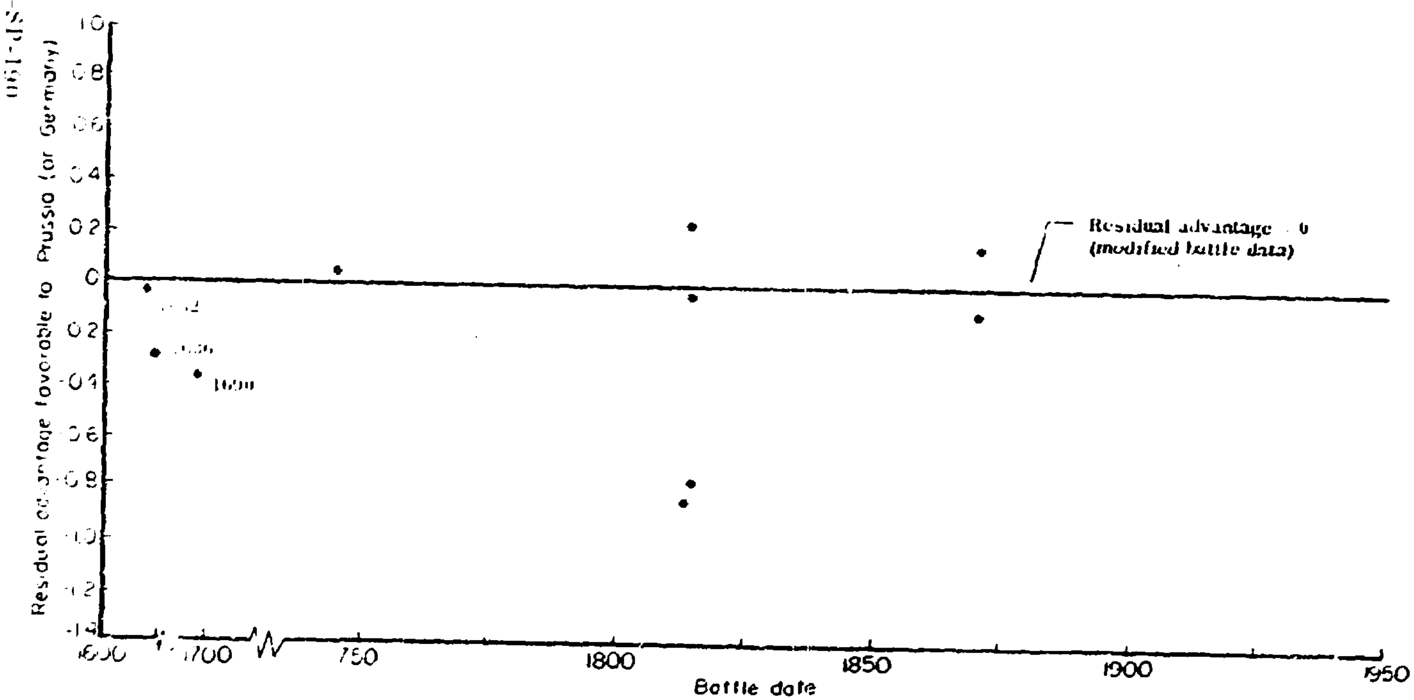


Figure 31. Linear scatter diagram of residual advantage favorable to Prussia (or Germany) against battle date for phase-2 data.

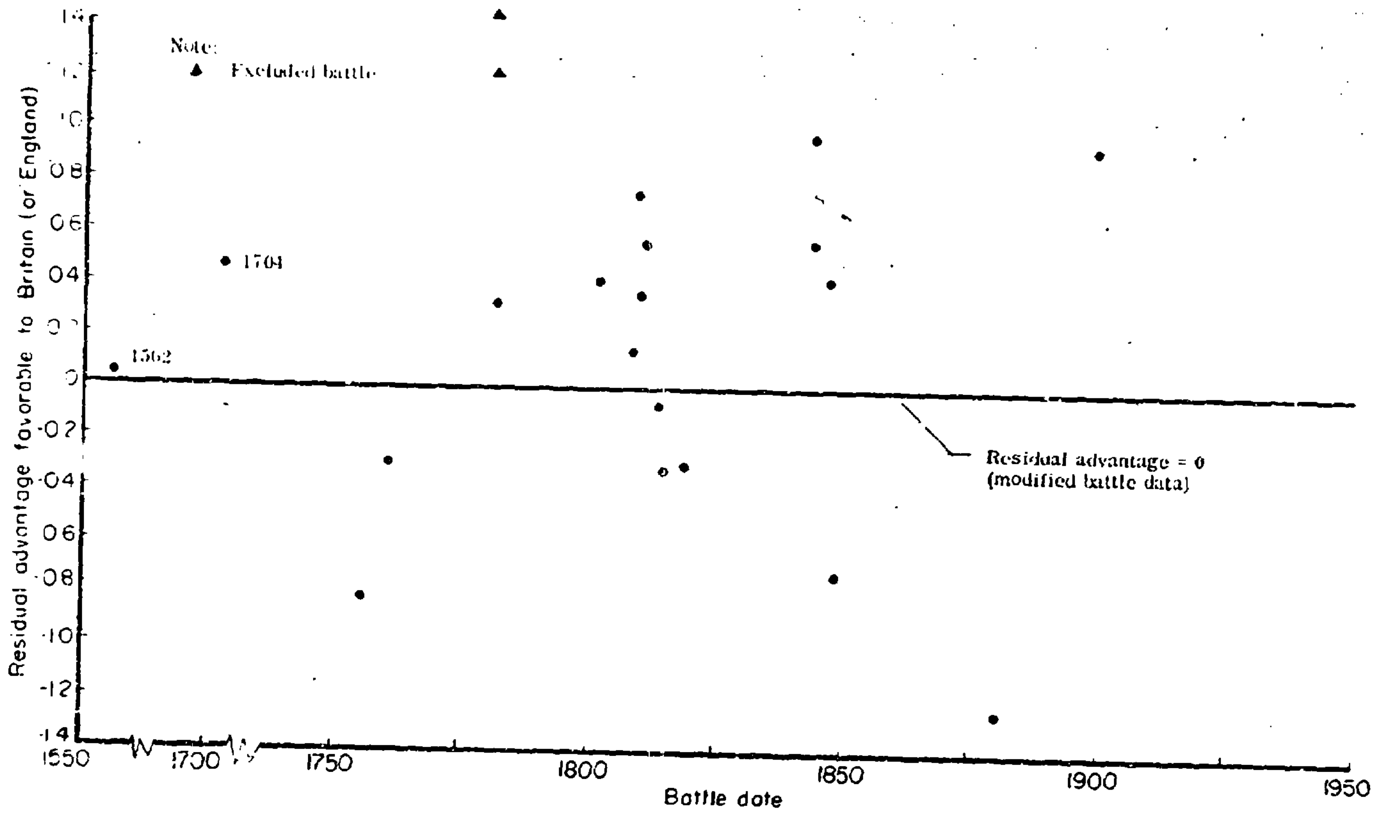


Figure 32. Linear scatter diagram of residual advantage favorable to Britain (or England) against battle date for phase-2 data.

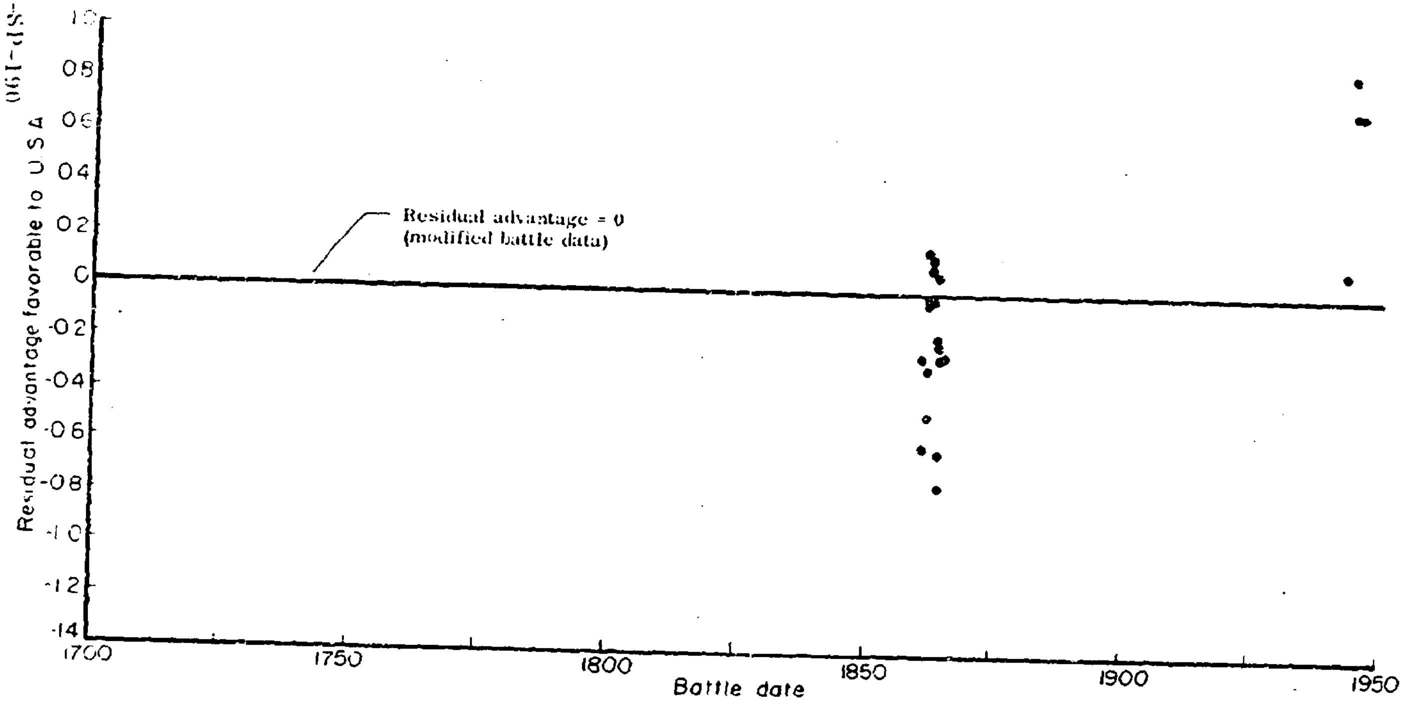


Figure 33. Linear scatter diagram of residual advantage favorable to United States against battle date for phase-2 data.

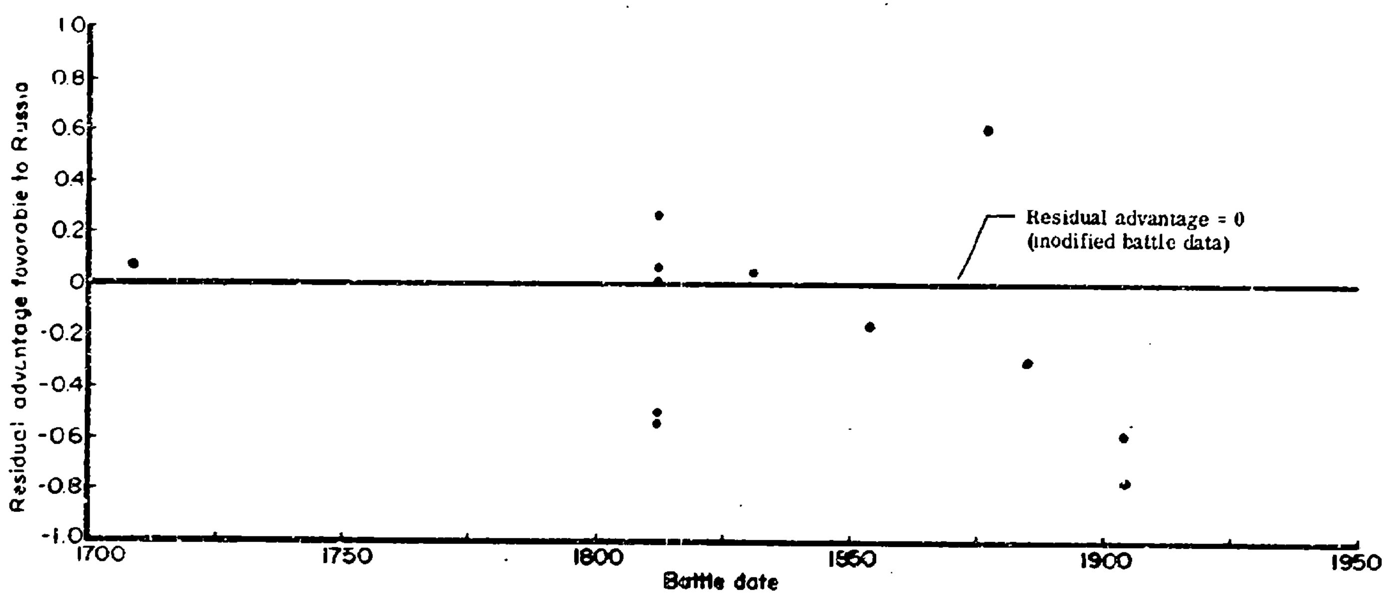


Figure 34. Linear scatter diagram of residual advantage favorable to Russia against battle date for phase-2 data.

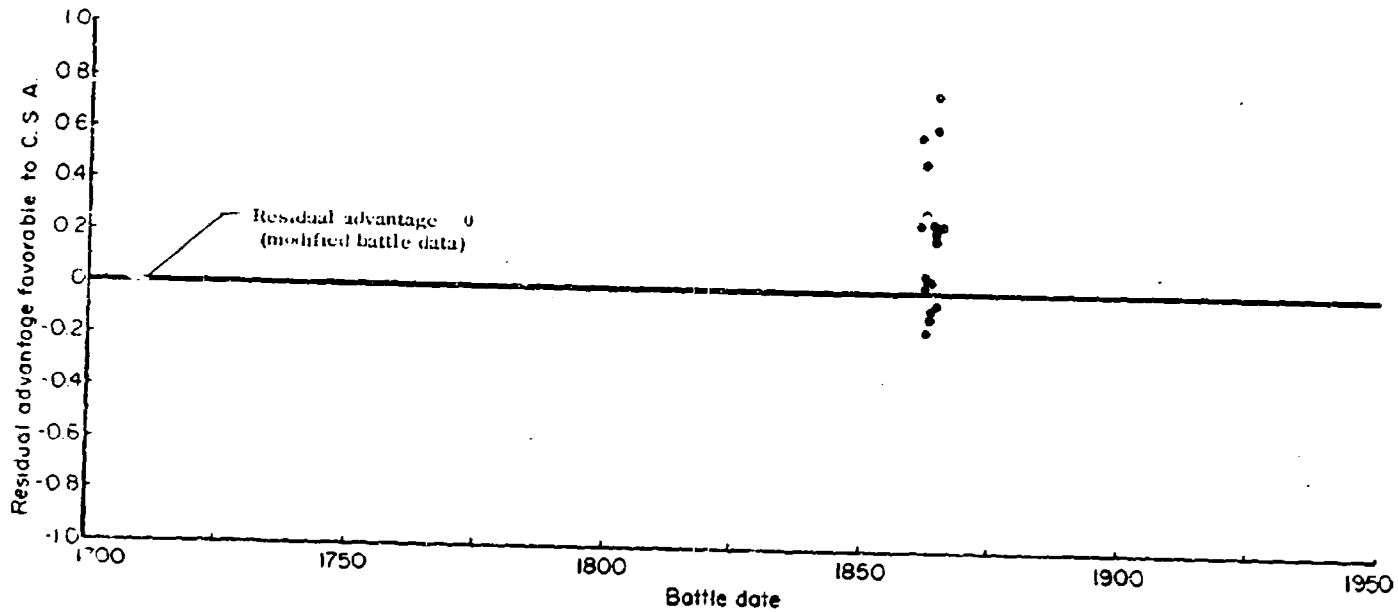


Figure 35. Linear scatter diagram of residual advantage favorable to C.S.A. against battle date for phase-2 data.

TABLE XXII

NUMBERS OF VICTORIES AND DEFEATS BY PARTICIPATING NATION AND BY SIDE

Participating Nation	Side on Which Nation Participated						Total Participation
	Attacking Side			Defending Side			
	Number of Victories	Number of Defeats	Total Participation on Attacking Side	Number of Victories	Number of Defeats	Total Participation on Defending Side	
France	12	8	20	6	5	11	31
United States of America	11	6	17	2	2	4	21
Britain (or England)	8	1	9	8	2	10	19
Confederate States of America	2	2	4	6	7	13	17
Russia	3	2	5	2	5	7	12
Prussia	2	3	5	1	4	5	10
India	1	2	3	0	5	5	8
Austria	0	2	2	1	4	5	7
Rome[a]	0	2	2	1	3	4	6
Japan	2	0	2	0	4	4	6
Sweden	1	1	2	2	1	3	5
Greece	1	0	1	1	0	1	2
Turkey	0	0	0	1	1	2	2
South Africa	1	0	1	0	1	1	2
Mexico	1	0	1	1	0	1	2
Switzerland	0	1	1	1	0	1	2
Canada	0	1	1	1	0	1	2
Eastern Roman Empire	1	0	1	0	0	0	1
Mongolia	1	0	1	0	0	0	1
Spain	0	0	0	0	1	1	1
Persia	0	1	1	0	0	0	1
Poland	0	0	0	0	1	1	1
Yugoslavia	0	1	1	0	0	0	1
Italy	0	1	1	0	0	0	1
Texas	0	0	0	0	1	1	1
Total	47	34	81	34	47	81	162

[a] Includes Western Roman Empire and Italy Roman Empire

TABLE XXIII

NUMBER AND FRACTION OF VICTORIES BY PARTICIPATING NATION

Participating Nation	Number of Victories	Total Participation	Victories as a Fraction of Total Participation	Approximate 95% Confidence Interval for Probability of Victory[a]
France	18	31	0.58	0.4-0.8
United States of America	13	21	0.62	0.4-0.8
Britain (or England)	16	19	0.84	0.6-1.0
Confederate States of America	8	17	0.47	0.3-0.7
Russia	5	12	0.42	0.2-0.7
Prussia (or Germany)	3	10	0.30	0.1-0.6
India	1	8	0.12	0.0-0.5
Austria	1	7	0.14	0.0-0.6
Rome	1	6	0.17	0.0-0.6
Japan	2	6	0.33	0.1-0.7
Sweden	3	5	0.60	0.2-0.9
Greece	2	2	1.00	0.2-1.0
Turkey	1	2	0.50	0.0-1.0
South Africa	1	2	0.50	0.0-1.0
Mexico	2	2	1.00	0.2-1.0
Switzerland	1	2	0.50	0.0-1.0
Canada	1	2	0.50	0.0-1.0
Eastern Roman Empire	1	1	1.00	0.0-1.0
Mongolia	1	1	1.00	0.0-1.0
Spain	0	1	0.00	0.0-1.0
Persia	0	1	0.00	0.0-1.0
Poland	0	1	0.00	0.0-1.0
Yugoslavia	0	1	0.00	0.0-1.0
Italy	0	1	0.00	0.0-1.0
Texas	0	1	0.00	0.0-1.0

[a]Estimates given are based on 95% confidence limits for the binomial probability parameter, computed by equating deserved fraction of victories to observed proportion of successes and total participation to sample size. Confidence intervals for each nation were independently determined and the interval limits were rounded to the nearest tenth prior to incorporation in the table.

COMPARISON OF PHASE-1 AND PHASE-2 RESULTS

In this section the results obtained in Reference 1 are compared with the results obtained in the foregoing for the modified data, and (to a lesser extent) with the entire set of phase-2 data. As each comparison is made, it is discussed in some detail so that conclusions may be drawn in the light of both phase-1 and phase-2 data. Additional discussion sections are included in order that additional background or other special material may be presented for consideration and evaluation by the reader.

Comparison of Data

Population and Sample

1. Sample Size

The total sample sizes for the phase-1 and the modified data are comparable (92 and 81 battles, respectively). However, the modified data have fewer battles with duration data (35 as compared to 82 of the phase-1 battles).

2. Distribution in Time

The modified data contains several battles which occurred prior to 1750, and thus extends the phase-1 data to earlier times.

3. Distribution in Space

The modified data contains battles from the East and African areas, though neither area is represented in the phase-1 data.

4. Distribution Among Countries

Five of the six most frequently participating nations, according to phase-1 data, also appear among the six most frequently participating nations according to the modified data.

5. Magnitude of the Sample Battles

In terms of total force, X, the modified data contain several small battles, though the phase-1 data did not. Moreover, the average total force for the modified data is about half that for the phase-1 data.

Comparative magnitudes in terms of total casualties, C, are similar to those stated above for total force, X.

The modified data exhibits about the same spread (or range of value) in total casualty fraction, F, values as phase-1 data. The average total casualty fraction value for the modified data is slightly larger than that for phase-1 data.

The modified data exhibits a greater spread, and a considerably higher average value, of battle duration, t, than phase-1 data.

Results for total manhours, M, are similar to those for battle duration.

6. Numerical Superiority

Force ratio, x_0 / y_0, values tend to be larger for modified data than for phase-1 data. Especially notable (e.g., in Fig. 3), is the much greater proportion of battles with high force ratios among the modified data than was the case for phase-1 data. The author is not surprised at this finding, since -in part- the phase-2 data were collected with the idea of extending phase-1 results to higher values of force ratio. As a result, a certain amount of extra effort was devoted to the collection of data on battles with large force ratio value.

7. Victorious Side

For neither phase-1 nor modified data is the proportion of attacker victories significantly different from $\frac{1}{2}$ (at the 5% level).

Comparison of Findings

Correspondence Between Parameters and Phenomena

1. Bitterness

Comparison of Figure 2 of this paper and Figure 2 of Reference 1 indicates that the phase-2 data scatters more widely about the curve $\epsilon = e^F - 1$ than does the phase-1. Some of the additional scatter exhibited by phase-2 data may be the result of inaccuracies in the historical record. However, certain of the data for World War II battles also depart from the trend, and the occurrence of fairly sizable errors in data for such recent battles seems unlikely.

We particularly note 15 battles which exhibit exceptionally wide deviation from the curve $\epsilon = e^F - 1$, namely, those battles numbered in Table I as 4, 5, 15, 16, 18, 22, 43, 45, 47, 74, 77, 80, 81, and 82. We define a set of battles called the "censored battle data" comprised of all phase-2 data except the above-listed 15 Exceptional battles[27].

2. Advantage

The advantage parameter follows the victor in a higher percentage (77.8%) of the modified battles than in the phase-1 data (73.9%).

A comparison of Table IV of this paper and Table IV of Reference 1 may suggest that the modified data has more exceptions to the general trend, which is that advantage follows the victor with greater fidelity for numerically large values of advantage. Nevertheless, a comparison of Table Va of this paper and Table Va of Reference 1 shows that, for advantage values numerically greater than 0.3, advantage for the modified

[27] See Appendix B for detailed descriptions of the various battle groupings used.

data follows the victor as frequently as for the phase-1 data (84.8% in each case). Also, chi-square values computed from the data of Table IIIa and Table Va are actually greater for the modified data than for the corresponding phase-1 data. Similar results would be obtained if data from all 83 phase-2 battles were used.

In sum, the phase-2 data strikingly confirm phase-1 findings with regard to the association between victory in battle and the Lanchester advantage parameter.

Behavior of Individual Parameters

1. Force Ratio

Figure 3 shows that the modified data contain a larger proportion of battles with very high force ratio values than the phase-1 data. As noted earlier, this may be a result of the special effort devoted to collection of phase-2 battles with very high force ratios. If some of these battles with a very high force ratio were deleted from the modified data, then the distribution of logarithmic force ratio for the modified data would be about the same as for phase-1 data. Consequently, the author feels that the phase-2 data tends to support rather than contradict phase-1 results regarding the distribution of logarithmic force ratio.

2. Advantage

Figure 5 shows that the modified data contains a larger proportion of battles with very low advantage values than the phase-1 data. Because of the negative correlation of advantage and force ratio, at least a part of this may be a result of the higher proportion of battles with very large force ratio values in the modified data. Deletion of the battles with very high force ratios would tend to bring the modified data in line with phase-1 [illegible]. In this case, as for the distribution of force ratio, the author feels that the modified data tend to support phase-1 results more than to contradict them.

The reader is cautioned, however, that formal tests have not been performed to rigorously establish the above arguments concerning the phase-2 distributions of force ratio and advantage. In part, this is because the author does not possess any criteria which would objectively select battles for deletion. Lacking such a criterion, the author prefers an uncomplicated subjective argument to an apparently sophisticated argument involving complicated procedures which might tend to suggest a higher degree of objectivity and rigor than is actually the case.

3. Bitterness

Despite an earlier finding to the effect that, for the modified data, logarithmic bitterness was (just barely) not normally distributed, we will in this section assume that it may be treated as being approximately normally distributed. Part of the justification for this is provided by Figure 6 which shows that ln c is in fact approximately normally distributed, and part is provided by the fact that the statistical procedures we will use are reasonably insensitive to moderate departures from strict normality.

We first wish to test whether the modified logarithmic bitterness data came from a population with a larger variance than the phase-1 data. Thus, we form the F-ratio of the sample variances[28], which amounts to about 1.21 with 91 and 80 degrees of freedom. On the hypothesis that the modified and phase-1 population variances are equal, a large F-ratio value could be expected to occur, by chance alone, more than 20 percent of the time (one-tailed test). Consequently, the data are consistent with the hypothesis that modified and phase-1 population of logarithmic bitterness, are equal. The estimated population variances, obtained by pooling the modified and phase-1 data[29], is 0.638, which corresponds to an estimated population standard deviation of 0.799.

---

[28] See, e.g., Reference 7.

[29] See Footnote 27.

Now we wish to test whether the modified data come from a population with a different average logarithmic bitterness than the phase-1 data. Following Reference 7, we compute a Student's t value of 1.01 at 171 degrees of freedom. On the hypothesis that modified and phase-1 populations have the same average value of logarithmic bitterness, a more extreme value of Student's t would be expected to occur by chance alone more than 30 percent of the time (two-tailed test). Thus, the data are consistent with the hypothesis that modified and phase-1 data come from populations with the same average logarithmic bitterness.

In sum, except for the marginally significant departure from strict normality of the modified logarithmic bitterness data, the data is consistent with the hypothesis that both modified and phase-1 logarithmic bitterness data represent independent samples from the same distribution. Appropriate pooled estimates of the mean and standard deviation of this common distribution are -2.099 and 0.799, respectively.

4. Surviving Fraction

Comparing Table IX of this paper and Table IX of Reference 1, we see that both modified and phase-1 data support the qualitative finding that the victorious side tends to have a larger surviving fraction than the defeated side. Indeed, the two tables are remarkably similar even from the quantitative viewpoint, save for the somewhat lower and more variable surviving defender fraction for case of attacker victory. (It is tempting to conjecture that this one discrepancy is connected with a higher proportion of battles with very large force ratios among the modified data.)

The same kind of similarity of modified and phase-1 data is evident on comparing Table X of this paper with Table X of Reference 1.

In sum, there is little evidence for believing that the modified and phase-1 battles come from radically different populations in regard to the behavior of surviving fraction data.

Relations Between Pairs of Parameters

1. Victory and Force Ratio

A comparison of Table XI of this paper and Table XIa of Reference 1 indicates that, for both modified and phase-1 data, force ratio follows the victor about 54.4 percent of the time. Modified, phase-1, and pooled (modified + phase-1) data are all consistent (at the 5% level of significance) with the hypothesis that force ratio follows victor 50 percent of the time.

Comparison of Table XII of this paper and Table XII of Reference 1 shows that modified, phase-1, and pooled (modified + phase-1) data are consistent (5% level of significance) with the hypothesis that force ratio and victory are independent.

Similar results obtain when battles with extreme force ratio values from modified and phase-1 data are compared with regard to the proportion of observed cases in which force ratio follows the victor.

These remarkable confirmations of phase-1 results leave little room for doubting the proposition that victory in battle is independent of numerical superiority[30].

2. Activity Ratio and Force Ratio

Figure 8 indicates that modified and phase-1 data are approximately the same with regard to the dependence of activity ratio on force ratio. The author prefers to deal with this and similar type questions via the dependence of advantage on force ratio, and this will be done in the following paragraphs.

---

[30] Combat models used for war gaming or other purposes should be considered in the light of these facts.

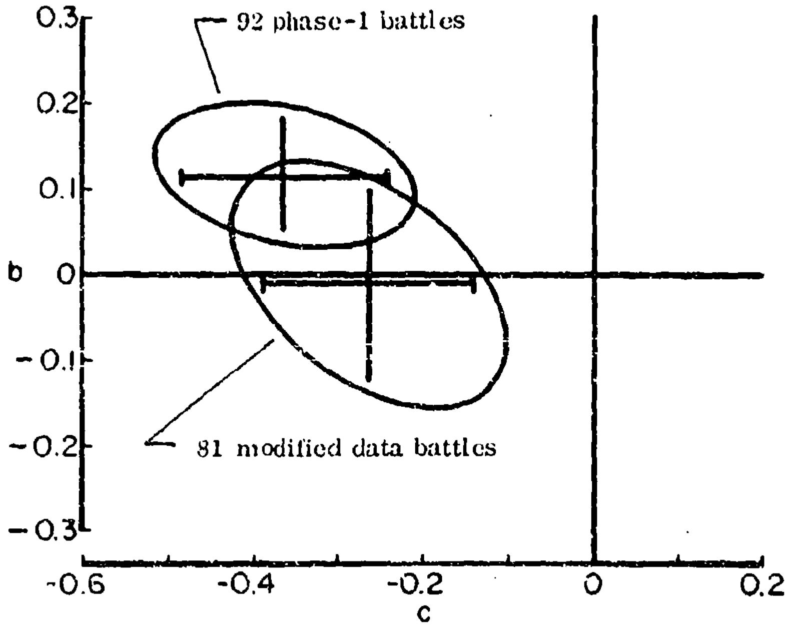


Figure 36. Ninety-five percent confidence intervals and regions for $\ln\mu = b + c\ln(x_0/y_0)$ based on phase-1 and modified data.

3. Defender Relative Advantage

a. Advantage and Force Ratio

Figure 7 suggests that modified and phase-1 data are similar in regard to the dependence of advantage on force ratio. It is desirable to have a somewhat more sophisticated method of comparing the two regression lines.

Unfortunately, the author has not been able to discover a rigorous exact statistical technique adequate to compare two sample regression lines under the conditions imposed by the data under consideration in this report. Consequently, comparison of phase-1 and other regression results is accomplished by the following approximate method:

(1) Estimated regression coefficients based on phase-1 data are assumed to be exact (i.e., free of sampling variability).

(2) Regression coefficients based on some other set of data are compared with the phase-1 regression coefficients to see if the 95 percent confidence region generated by the other data includes the phase-1 regression point or not[31].

Figure 36 shows how the approximate method is applied to compare phase-1 and modified data with respect to the regression of advantage on logarithmic force ratio. Note that the ordinate and abcissa of Figure 36 represent the intercept and slope, respectively, of the population regression line. Confidence intervals and region, based on phase-1 and modified data regres[illegible] coefficients b and c, are plotted. These confidence intervals and regi[illegible] e centered on the maximum likelihood estimates plot as points on axes [illegible]senting population values of th[illegible] egression coefficients (such as Fi[illegible]) as a "phase diagram."[32]

[31]Se[illegible] Re[illegible] 3 for a discussion of confidence regions and the method of u[illegible] [illegible]m[illegible] [illegible]pare sample regression lines with a priori regression coefficients.

[32]The reader is cautioned that this may not correspond to standard terminology.

Thus we refer to Figure 36 as the phase diagram showing confidence intervals and regions for the phase-1 and modified data regressions of advantage on force ratio.

In accord with the criteria for comparing regression lines outlined above, the phase-1 confidence region is considered as being concentrated at the phase-1 maximum likelihood point and the phase diagram is examined to see whether the modified data confidence region includes that point or not. In Figure 36, the modified data confidence region does not include the phase-1 maximum likelihood point. This indicates that the modified data is consistent with the phase-1 data with regard to the regression line of advantage on logarithmic force ratio[33] .

---

[33] Note that since we are using 95 percent confidence regions throughout, and since we are also neglecting sampling variability in the phase-1 data, we in fact have consistency either at the 5 percent level or at a higher level. That is, a worse agreement would be expected to occur, through the action of chance alone, at least 5 percent of the time. In statistical terminology, the test criterion outlined above will reject the null hypothesis (of no difference in phase-1 and other data population regression lines) at least 5 percent of the time by chance alone, assuming that the null hypothesis is true, that is, the probability of a Type 1 error is 5 percent or greater. Thus the proposed criteria for comparing regression lines tends to be on the conservative side in the sense that there are some circumstances for which it will lead to rejection of the null hypothesis while in the same circumstances a more exact treatment would lead to acceptance of the null hypothesis, and whenever the exact treatment leads to rejection of the null hypothesis the approximate criteria will also lead to rejection of the null hypothesis.

These statistical assertions are intuitively obvious when it is remembered that the approximate criteria correspond to the exact treatment for a test which assumes phase-1 results to be given a priori, that is, without allowance for sample variability. The exact treatment of this problem would make appropriate allowances for the sampling variability of phase-1 maximum likelihood points.

Although we have argued in the preceeding paragraph that the population regression line of advantage on logarithmic force ratio is the same for modified as for phase-1 data, a comparison of Table XIII and Table XIII of Reference 1 indicates that the variability of residual advantage about the regression line is greater for the modified than for the phase-1 data. In fact, the variance F-ratio of modified to phase-1 residual advantage data amounts to about 2.40 with 79 and 90 degrees of freedom, a value which would be exceeded by chance alone less than 0.5 percent of the time (one-tailed test). Thus, the data indicate that the modified data tend to scatter more widely about the regression line than do phase-1 data.

In an effort to identify, if possible, the source of the increased variability observed in the modified data, the phase-2 data were arranged to form various battle groupings. Five such battle groupings are analyzed in the following paragraphs. They are designed as follows:

Battle Grouping	Designation	Approximate Description[34]
1.	Phase-2	83 battles as listed in Table I
2.	Modified	81 battles (battles number 21 and 22 are deleted from phase-2 grouping)
3.	Censored	68 battles (15 battles exhibiting wide departure from $\epsilon = e^F-1$, are deleted from phase-2 grouping)
4.	Harbottle	56 battles (include those and only those phase-2 battles for which Reference 2 was used as a source of quantitative data)
5.	Non-Harbottle	27 battles (56 Harbottle battles are deleted from phase-2 grouping)

[34] See Appendix B for a detailed description of the various battle groupings employed.

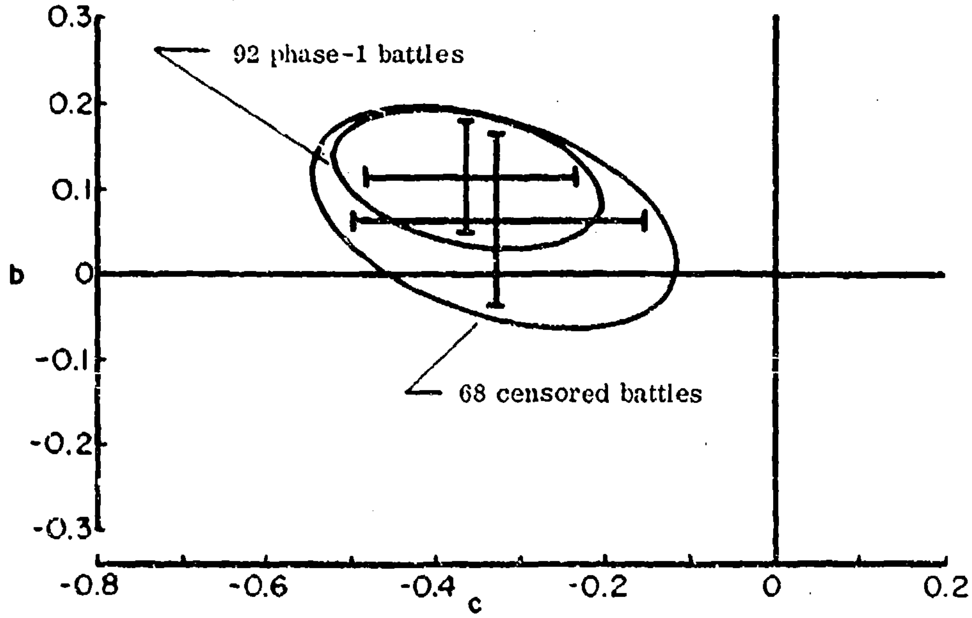
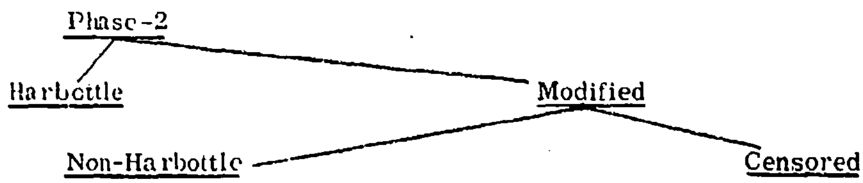


Figure 37. Ninety-five percent confidence intervals and regions for $\ln\mu = b + c \ln(x_0/y_0)$ based on phase-1 and censored battle data.

Some of the relations among these five battle groupings are readily grasped through inspection of the following diagram, where a slanting line connecting two battle groupings means that the lower-placed grouping is included in the higher-placed grouping. (Note that battle groupings on the same horizontal level are not necessarily mutually exclusive.)



The regression lines of advantage on logarithmic force ratio, calculated on the basis of these battle groupings, are compared to the corresponding phase-1 regression line using the phase diagram technique. Figure 37 shows the comparison of phase-1 and censored data, and indicates a higher degree of agreement than that shown in Figure 36. Figure 38 shows the comparison of phase-1 data with phase-2 Harbottle, and Non-Harbottle data[35]. Figure 38 shows that, although the phase-2 data is not consistent with phase-1 data (using the approximate criteria enunciated earlier in this section), the phase-2 data still lead to a slope, c , of advantage versus logarithmic force ratio which is significantly different from zero (5% level). Such is not the case for the Harbottle data. For clarity, confidence regions for the Harbottle and Non-Harbottle data are not shown in Figure 38. Confidence regions for these battle groupings would be elliptical in shape and would be related to the corresponding confidence intervals in a manner like that exhibited by other battle groupings. Specifically, the 95-percent confidence region for the Non-Harbottle data are consistent with phase-1 data with respect to the approximate comparison criteria given above, while Harbottle data would not be consistent with phase-2 data.

---

[35] The Harbottle and Non-Harbottle groupings are, of course, mutually exclusive and exhaustive with respect to phase-2 data.

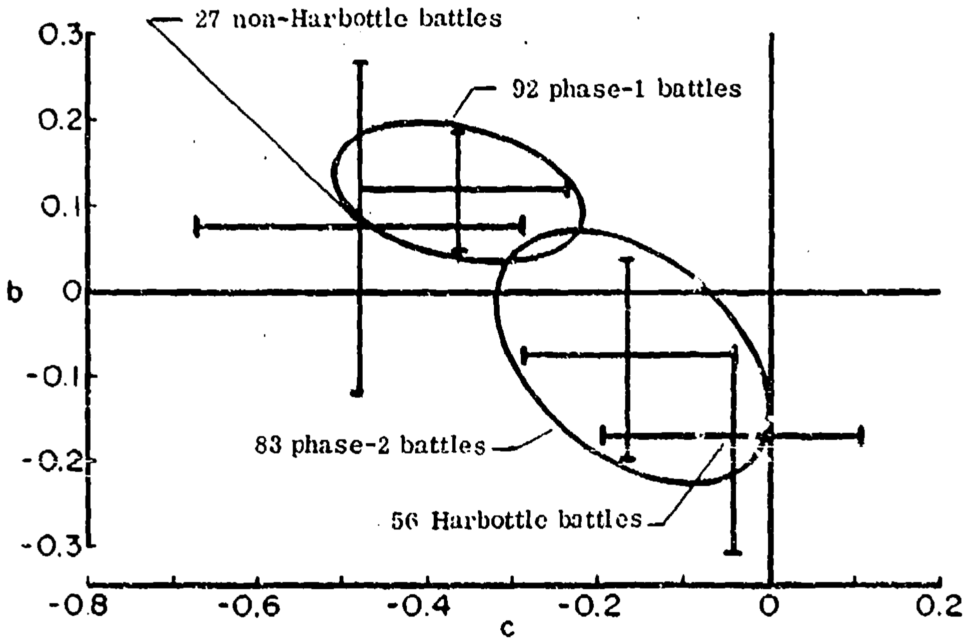


Figure 38. Ninety-five percent confidence intervals and/or regions for $\ln \mu = b + c \ln(x_0 \, y_0)$ based on phase-1 data, phase-2 data, non-Harbottle data, and on Harbottle data.

Consideration of Figures 26, 37, and 38 in the light of relations among the battle groupings displayed in the diagram given above suggests that the greater the proportion of Harbottle data, the worse the degree of agreement with phase-1 data tends to be. This suggests that the Harbottle data may include a higher proportion of battles with erroneous data than for the Non-Harbottle data.

Table XXIV tabulates, for each battle grouping, the residual variability (variability of the data about that grouping's regression line) together with F-ratios and approximate (one-tailed) significance levels. Of the five battle groupings considered, only the Non-Harbottle grouping is consistent with phase-1 data with respect to residual variability. Thus, the data of Table XXIV also show that battle groupings with higher proportion of Harbottle data tend to exhibit poorer agreement with a small proportion of Harbottle data.

In the light of the facts outlined above, the author feels impelled to conclude that the Harbottle data, derived from Reference 2, are afflicted with large and/or more numerous errors than is the case for data obtained from other sources consulted[36]. Although this conclusion is based largely on

[36] This finding, while perhaps suggesting that less use ought to have been made of the Harbottle data, does illustrate a possible application of the phase-1 and phase-2 findings to certain aspects of historical criticism.

TABLE XXIV

COMPARISON OF RESIDUAL VARIABILITY FOR THE REGRESSION OF ADVANTAGE ON LOGARITHMIC FORCE RATIO FOR VARIOUS BATTLE GROUPINGS

Battle grouping	Residual variability, $s^2_{\ln\mu \mid \ln x_0/y_0}$	Number of data points	Variance F-ratio[a]		Significance level[b] (one-tailed test)
			Sample value	Degrees of freedom	
Harbottle	0.254	56	2.886	54/90	***
Phase-2	0.241	83	2.739	81/90	***
Modified	0.211	81	2.398	79/90	***
Censored	0.151	68	1.716	66/90	**
Non-Harbottle	0.140	27	1.591	25/90	*
Phase-1	0.088	92	1.000	-----	----

[a] Phase-1 estimated residual variance in the denominator

[b] *** = Significant at 0.5% level or better

** = Significant at 1% level, but not at 0.5% level

* = Significant at 10% level, but not at 5% level

TABLE XXV

COMPARISON OF THE DISTRIBUTION OF MODIFIED HARBOTTLE AND NON-HARBOTTLE BATTLES WITH RESPECT TO BATTLE DATE

Time Interval	Harbottle Battles		Non-Harbottle Battles		Modified Battles	
	Number	Percent	Number	Percent	Number	Percent
500 B. C. - 1499	5	9	0	0	5	6
1500 - 1599	0	0	3	11	3	4
1600 - 1699	3	6	3	11	6	7
1700 - 1799	6	11	2	7	8	10
1800 - 1849	23	43	3	11	26	32
1850 - 1899	16	30	11	41	27	33
1900 - 1949	1	2	5	19	6	7
Total	54	101	27	100	81	99

on evidence internal to the phase-2 samples of battles--and there are many objections to such a procedure from the standpoint of strict rigor--the author finds the argument highly convincing [37].

In sum, the author believes that the phase-2 data does not contradict phase-1 findings regarding the relation between advantage and logarithmic force ratio. At the same time, it must be admitted that phase-2 data do not support the phase-1 findings as strongly as the author hoped they would. Nevertheless, the author is convinced that the phase-1 findings are valid, and that data from Harbottle's Dictionary (Reference 2) should be accepted only when independent sources confirm its accuracy.

b. Advantage and Bitterness

It has been argued in an earlier section that the correlation of logarithmic bitterness on advantage summarized in Table XV can be considered as the result of a purely chance effect. Comparison of Table XXVI with Table XV of Reference 1 shows that the Non-Harbottle data are consistent with phase-1 data in regard to the regression of logarithmic bitterness on advantage.

---

[37] The reader is also referred to the discussion in Appendix B to Reference 1 of (i) Harbottle's Dictionary and (ii) a limited comparison of Harbottle data with that from other sources.

One of the objections the reader might raise against the argument that Harbottle data are inaccurate is that the Harbottle data tend to contain a higher proportion of battles of very early date than the Non-Harbottle data (i.e., that early battle data and use of Harbottle as a data source are confounded factors in the statistical sense). Actually, this objection is not sound as an inspection of Table XXV will show. Note that the distribution of both Harbottle and Non-Harbottle battles is essentially similar to that for modified battles with the exception of the two half-centuries 1800-1849 and 1900-1949. Thus, Harbottle data do not, in fact, contain a much higher proportion of battles of very early date than the Non-Harbottle data.

Another objection that might be raised by the reader is to assert that the Harbottle data are "right" and the Non-Harbottle and phase-1 are "wrong". This objection cannot be satisfactorily resolved on the basis of internal evidence alone.

TABLE XXVI

RESULTS OF REGRESSION OF LOGARITHMIC BITTERNESS, $\ln \epsilon$, ON ADVANTAGE, $\ln \mu$, FOR NON-HARBOTTLE DATA

Regression line: $\ln \epsilon = b + c \ln \mu$

Number of data points: 27

Estimated value of regression coefficients ±95% confidence limits:

$b = -1.897 \pm 0.388$

$c = -0.361 \pm 0.694$

Standard error of estimate: $s_{\ln \epsilon \mid \ln \mu} = 0.897$

Correlation coefficient: $r = -0.209$

Sample mean:

of $\ln \epsilon = -1.816$

of $\ln \mu = -0.226$

Sample variance:

of $\ln \epsilon = 0.809$

of $\ln \mu \approx 0.271$

In view of the doubt cast on the validity of Harbottle data by the previous section, this finding tends to confirm the propriety of rejecting the correlation of logarithmic bitterness on advantage exhibited by modified data as an artifact of the data.

c. Advantage and Battle Duration

Comparison of Table XVI and Table XVI of Reference 1 shows that modified and phase-1 data are in agreement with regard to the absence of correlation between advantage and logarithmic duration. The standard error of estimate, for modified data is significantly larger than for phase-1 data (F-ratio of 2.04 at 32/80 degrees of freedom, which is significant at better than the 1% level, using the one-tailed test), but this could be due in part to Harbottle data inaccuracy.

d. Residual Advantage and Various Other Parameters

(1) Bitterness. Figure 14 and Figure 14 of Reference 1 both indicate that there is no correlation between residual advantage and logarithmic bitterness.

(2) Total Force. Figure 15 and Figure 15 of Reference 1 both indicate that there is no correlation between residual advantage and logarithmic total force.

(3) Total Casualties. Figure 16 and Figure 16 of Reference 1 both indicate that there is no correlation between residual advantage and logarithmic total casualties.

e. Residual Advantage and Battle Date

Figure 17 and Figure 17 of Reference 1 both indicate that there is no consistent trend of residual advantage with battle date.

f. Narrative Accounts of Battles with Extreme Residual Advantage Values

Comparison of Appendix C with Appendix C of Reference 1 suggests that, for both modified and phase-1 data, the achievement by the attacker of surprise may result in extreme negative residual advantage values, and that failure of the attacker to achieve surprise may result in extreme positive residual advantage values.

The reader is cautioned that the small number of narrative accounts together with their brevity and lack of specific detail, do not permit firm conclusions to be drawn. Thus, the observations of the preceeding paragraph must be counted for the most part as interesting speculations.

g. Residual Advantage and Victory

Comparison of Table XVIIa and Table XVIIa of Reference 1 shows that residual advantage follows the victor 77.8 percent of the time for the modified data and 70.6 percent of the time for all phase-1 data. Both modified and phase-1 data indicate that victory and residual advantage are significantly (at much better than the 0.5% level) associated.

Comparison of Table XVIII and Table XVIII of Reference 1 indicates that residual advantage follows the victor with greater fidelity for numerically large values of residual advantage than for numerically small values in both the modified and the phase-1 data.

Comparison of Table XIXa and Table XIXa of Reference 1 shows that, for residual advantage values numerically greater than 0.2, residual advantage follows the victor 84.3 percent of the time for modified data and 78.0 percent of the time for phase-1 data. Both modified and phase-1 data indicate that victory and residual advantage are significantly (at much better than the 0.5% level) associated.

In sum, the modified data strikingly confirm the phase-1 findings with regard to the association of victory in battle and the residual advantage parameter.

4. The Effects of Battle Date

a. Force Ratio and Battle Date

Comparison of Figure 18 and Figure 18 of Reference 1 indicates that there is no consistent trend of logarithmic force ratio with battle data for either modified or phase-1 data.

b. Duration and Battle Date

Comparison of Figure 19 and Figure 19 of Reference 1 suggests that, for both modified and phase-1 data, battle duration has tended to increase on the average, since about 1750.

c. Bitterness and Battle Date

Figure 20 of Reference 1 was judged on an informal basis to exhibit no consistent trend of bitterness with battle date. In an earlier section of this paper, it has been suggested that Figure 20 hints at a very gradual and long-term decline in bitterness with battle date. Comparison of Figure 20 and Figure 20 of Reference 1 indicates that, for both phase-1 and modified data, World War II battles may have been more bitter than most battles which occurred since 1750.

d. Intensity and Battle Date

Comparison of Figure 21 and Figure 21 of Reference 1 suggests that, for both modified and phase-1 data, there has been a tendency to decline with increasing battle date since 1750. The World War II battles appear to fall more nearly in line with the general trend for the modified data than they did for the phase-1 data.

5. Bitterness, Intensity, and Duration

A careful comparison of Figure 22 with Figure 22 of Reference 1, and of Table XX with Table XX of Reference 1 may leave some doubt as to whether the modified data are consistent with the phase-1 data in regard to the relation between logarithmic bitterness and logarithmic duration. We will return to this question when we compare modified and phase-1 data with regard to the relation between logarithmic intensity and logarithmic duration, where it will be shown that the two sets of data are consistent.

a. Residual Logarithmic Bitterness

Despite the finding of Reference 1 to the effect that, for phase-1 data, residual logarithmic bitterness was (at a marginal level of significance) not normally distributed, we will, in this section, assume that it may be treated as being approximately normally distributed. Part of the justification for this is provided by Figure 23 of Reference 1 which shows that residual logarithmic bitterness is in fact approximately normally distributed, and part is provided by the fact that the statistical procedures we will use are reasonably insensitive to moderate departures from strict normality.

We wish to test whether the modified data come from a population with a larger variance than the phase-1 data. The F-ratio of the sample variances amounts to about 1.43 with 33/90 degrees of freedom. A large F-ratio value could be expected, by chance alone, about 10 percent of the time (one-tailed test), so the data are consistent with the hypothesis that modified and phase-1 population variances are equal.

Comparison of Figure 24 and Figure 24 of Reference 1 indicates that residual logarithmic bitterness does not exhibit a consistent trend with battle date for either modified or for phase-1 data.

b. Intensity and Duration

Figure 25 shows the phase-1 and modified data regression lines of logarithmic intensity on logarithmic duration. Figure 39 shows a phase diagram for the regression line confidence intervals and regions. Based on the criteria described in a previous section for comparing regression lines, the modified data is consistent with phase-1 results with regard to the relation between intensity and duration[38].

6. Total Casualties and Total Force

Comparison of Figure 26 with Figure 26 of Reference 1 indicates that the line $C = 15X$ provides a reasonable fit to both modified and phase-1 data.

7. Force Ratio and Total Force

Comparison of Figure 27 and Figure 27 of Reference 1 indicates that neither modified nor phase-1 data exhibit any consistent trend of force ratio with total force. Average logarithmic force ratio tends to be somewhat larger for the modified data than for the phase-1 data, but this is probably due to the unusually large proportion of battles with very large force ratio values in the modified data.

---

[38] In an earlier section dealing with a comparison of phase-1 and phase-2 data findings with regard to the relation between advantage and logrithmic force ratio, it was argued that battle groupings with a high proportion of Harbottle data sometimes tend to exhibit poor agreement with phase-1 results. Remembering this finding, it is interesting to note that only a relatively small proportion of the modified battles for which duration data are available use Reference 2 as a source of data (11 out of 35, or less than a third of these battles use Harbottle's Dictionary as a source). This fact may help to explain how errors in the Harbottle data could significantly affect the relations between advantage and force ratio without significantly influencing the relation between intensity and duration. (For comparison, $\frac{2}{3}$ of the modified data battles are Harbottle battles.)

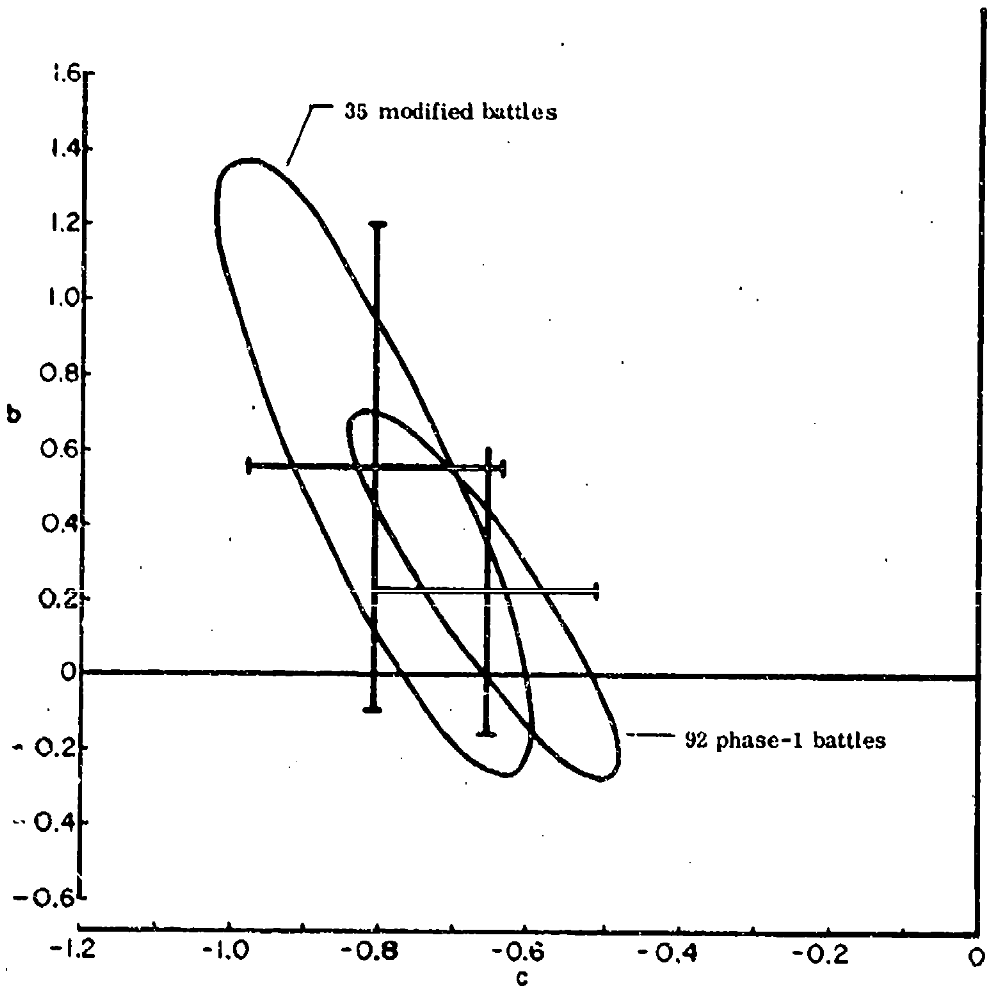


Figure 39. Ninety-five percent confidence limits and regions for $\ln\lambda = b + c\ \ln t$ based on 92 phase-1 battles and on 35 modified battles

8. Duration, Total Force, and Force Ratio

Comparison of Figure 28 with Figure 28 of Reference 1 indicates that neither modified nor phase-1 data exhibit any obvious correlation of duration with total force.

Comparison of Figure 29 with Figure 29 of Reference 1 indicates that neither modified nor phase-1 data exhibit any obvious correlation of duration with force ratio.

9. Participating Nation and Victory

Comparison of Figure 30 through Figure 35 with Figure 30 through Figure 34 of Reference 1 indicates that, whether the modified or phase-1 data is considered, no one nation has consistently had a favorable residual advantage.

Comparison of Table XXIII with Table XXIII of Reference 1 indicates that, whether the modified or phase-1 data is considered, no one nation seems to have consistently been tactically superior to its opponents.

Recapitulation of Findings

In the following we attempt to note as briefly as possible, the areas of agreement (to within inherent data variability) or disagreement between the findings of this paper and those of Reference 1. In cases of disagreement, the author will briefly state his current convictions, based on available information, or call attention to unresolved issues.

1. The Lanchester bitterness parameter, ϵ, is a reasonably good index of total casualty fraction. The modified data exhibit some cases of wide deviation from the trend curve $\epsilon = e^F - 1$, so the issue has not been resolved in entirely satisfactory fashion.

2. The Lanchester advantage parameter, ln μ , is closely associated with probability of victory. Agreement.

3. The other Lanchester parameters are valid indices of the real-world phenomena identified by the names of the respective parameters. Unresolved. Author is convinced that the assertion is correct.

4. Logarithmic force ratio, ln x_0/y_0 , is approximately normally distributed with mean 0.2 and standard deviation 0.5. Disagreement. Author is convinced that the assertion is correct in at least a rule-of-thumb sense[39] .

5. Advantage, ln μ , is approximately normally distributed with mean 0.0 and standard deviation 0.4. Disagreement. Author is convinced that the assertion is correct in at least a rule-of-thumb sense[40] .

6. Logarithmic bitterness, ln ϵ , is approximately normally distributed with mean -2.1 and standard deviation 0.8. Agreement[41] .

7. There is a tendency for the victorious side to have a larger and less variable surviving fraction than the defeated side. Agreement.

8. For the range of values represented in the data, victory is, at best, only weakly related to force ratio. Agreement.

9. Logarithmic activity ratio, ln D/A , tends to increase with <u>increasing</u> logarithmic force ratio, ln x_0/y_0. Agreement.

---

[39] Numerical values are based primarily on phase-1 data.

[40] See footnote 37.

[41] Numerical values are based on pooled phase-1 and modified data.

10. Advantage, $\ln \mu$, tends to decrease with increasing force ratio, $\ln x_0 / y_0$, approximately according to the equation (see Table XXVII)[42]:

$$\ln \mu = 0.1 \ln x_0 / y_0$$

There is agreement between the modified and phase-1 data with regard to the population regression line, but not with regard to the variability of residual advantage. The author believes that the disagreement originates primarily with inaccuracies in the Harbottle data.

11. Advantage, $\ln \mu$, and logarithmic bitterness, $\ln \epsilon$, are uncorrelated. Disagreement. The author believes that the significant correlations exhibited by the modified data originates primarily in Harbottle data errors.

12. Advantage, $\ln \mu$, and logarithmic duration, $\ln t$, are uncorrelated. Agreement.

13. Residual advantage and logarithmic bitterness, $\ln \epsilon$, appear to be uncorrelated. Agreement.

14. Residual advantage and logarithmic total force, $\ln X$, appear to be uncorrelated. Agreement.

15. Residual advantage and logarithmic total casualties, $\ln C$, appear to be uncorrelated. Agreement.

16. Residual advantage and battle date appear to be uncorrelated. Modified and phase-1 data agree, except for some phase-1 battles which occurred during the period 1757 and 1760. The author believes that this exception is more accidental than real.

17. Data inaccuracies do not account for all of the observed variability in residual advantage[43]. Agreement.

---

[42] See footnote 39.

[43] Based in part on the between-source data variability estimates developed in Appendix B of Reference 1.

TABLE XXVII

RESULTS OF REGRESSION OF ADVANTAGE, $\ln \mu$, ON LOGARITHMIC FORCE RATIO, $\ln x_0/y_0$, FOR POOLED (MODIFIED + PHASE-I) DATA

Regression line: $\ln \mu = b + c \ln x_0/y_0$

Number of data points: 173

Estimated value of regression coefficients ±95% confidence limits:

$b = 0.061 \pm 0.064$

$c = -0.311 \pm 0.088$

Standard error of estimate: $s_{\ln \mu \mid \ln x_0/y_0} = 0.384$

Correlation coefficient: $r = -0.480$

Sample mean:

of $\ln \mu = -0.025$

of $\ln x_0/y_0 = 0.276$

Sample variance:

of $\ln \mu$ 0.190

of $\ln x_0/y_0 = 0.452$

18. Comparison of narrative accounts of battles with unusually large and unusually small advantage values suggests that successfully achieving surprise tends to be highly advantageous to the attacker, and that failure of the attacker to achieve surprise tends to be highly advantageous to the defender. Issue is still unresolved. More, and more detailed, comparisons of narrative accounts than could be attempted in this study are necessary to resolve this issue.

19. The principal determinants of victory in battle are contained in the residual advantage parameter. Agreement.

20. Logarithmic force ratio, $\ln x_0/y_0$, and the battle date appear to be uncorrelated. Agreement.

21. Logarithmic duration, $\ln t$, has exhibited a tendency to increase since about 1750. Agreement.

22. Logarithmic bitterness, $\ln \epsilon$, seems to exhibit a very gradual long-term trend toward bitterness values. World War II Pacific Ocean island campaigns tend to be more bitter than the average battle. There is approximate agreement between modified and phase-1 data. Additional data, particularly for battles of very early date, are required for adequate resolution of this issue.

23. Logarithmic intensity, $\ln \lambda$, appears to have gradually declined since about 1750. The activity parameters A and D generally follow the same trend. Agreement.

24. Logarithmic bitterness, $\ln \epsilon$, tends to increase with logarithmic duration, $\ln t$. Agreement.

25. Data inaccuracies large enough to account for all of the observed variability in residual logarithmic bitterness[44]. Agreement.

---

[44]See footnote 41.

26. Residual logarithmic bitterness is approximately normally distributed with mean 0.0 and standard deviation 0.8. Disagreement. The author is convinced the assertion is valid, at least in a rule-of-thumb sense.

27. Residual logarithmic bitterness and battle date appear to be uncorrelated. Modified and phase-1 data agree, except for some phase-1 battles which occurred during the decade 1840-1849. The author believes that this exception is due to data inaccuracies.

28. Logarithmic intensity, $\ln \lambda$, tends to decrease with increasing logarithmic duration, $\ln t$, approximately according to the equation (see Table XXVIII)[45] :

$$\ln \lambda = 0.4 - 0.75 \ln t$$

There is agreement between the modified and phase-1 data both with regard to the population regression line and with regard to the variability about the regression line.

29. Logarithmic total casualties, $\ln C$, tend to increase with logarithmic total force, $\ln X$, approximately in accord with the equation:

$$C = 0.15 X.$$

Modified and phase-1 data appear to agree with regard to the population regression line, but no formal analysis has been attempted. The author is quite confident that the assertion is valid.

30. Logarithmic force ratio, $\ln x_0/y_0$, appears to be uncorrelated with logarithmic total force, $\ln X$. Agreement.

31. It appears that no nation has consistently been tactically superior to its opponents. Agreement.

---

[45] See footnote 39.

TABLE XXVIII

RESULTS OF REGRESSION OF LOGARITHMIC INTENSITY, $\ln \lambda$, ON LOGARITHMIC DURATION, $\ln t$, FOR POOLED (MODIFIED - PHASE-I) DATA

Regression line: $\ln \lambda = b + c \ln t$

Number of data points: 117

Estimated value of regression coefficients ±95% confidence limits:

$b = 0.408 \pm 0.304$

$c = 0.746 \pm 0.100$

Standard error of estimate: $s_{\ln \lambda | \ln t} = 0.790$

Correlation coefficient: $r = -0.809$

Sample mean:

of $\ln \lambda = -1.589$

of $\ln t = 2.678$

Sample variance:

of $\ln \lambda = 1.791$

of $\ln t = 2.108$

In general, the analysis of modified or phase-2 data strongly supports phase-1 results. Important major areas of disagreement are:

1. The relation between bitterness, ϵ , and total casualty fraction, F.

2. The amount of variability about the regression line of advantage, $\ln \mu$, on logarithmic force ratio, $\ln x_0/y_0$.

3. The independence of advantage, $\ln \mu$, and logarithmic bitterness, $\ln \epsilon$

DISCUSSION

Future Work

It is of course admitted that this paper does not contain all the information we would like to have. Improvements achievable by further study are discussed below.

1. More precise specification of regression coefficients and other descriptive quantities as well as clarification of some of the remaining doubtful issues can readily be obtained from the analysis of additional data samples.

2. A fuller understanding of the empirical findings depends upon the construction of a theoretical framework adequate to explain the principal experimental facts.

3. Additional descriptive work is needed on the factors which give rise to the variability of residual advantage. At present, the only available approach seems to be in the analysis of a large number of narrative accounts of battles with the object of separating out those factors which are common to many battles with extreme residual advantage values. Of course, the results of such a study would have to be verified by testing them against additional data.

Unfortunately, a very considerable effort will be necessary to collect a sufficiently large sample of narrative accounts, especially since these must be quite detailed in order to ensure that the narratives include all of the key factors. Thus, the author estimates that an effort approaching a level of two man-years would be required to properly carry out a preliminary study of factors affecting residual advantage, and would probably demand the services of a highly competent military historian working in conjunction with an able statistician. This kind of activity could probably be more efficiently performed if an adequate theoretical framework is available to guide the work. Otherwise, while the basic idea is simple enough, the amount of work required becomes onerous.

4. If a study of the narrative accounts should be pursued, duplication of effort can be minimized by establishing a parallel study of the factors affecting intensity and/or duration of battle since these studies can make use of the same narrative accounts.

Applications

In Reference 1 it was suggested that the historical findings could be applied to assist in judging the validity of a ground combat war game and in analyzing game results. An example of how to do this is presented in Reference 10.

APPENDIX A

NARRATIVE ACCOUNTS OF TWO EXCLUDED BATTLES

Battle of Porto Novo, 1781, First Mysore War

Narrative from Reference 2

"Fought July 1, 1781, between 8,500 British troops, under Sir Eyre Coote, and about 65,000 Mysores, under Hyder Ali. Hyder occupied a strongly entrenched camp, blocking the British advance upon Cuddalore. Here he was attacked by Coote, and after a day's hard fighting the position was stormed, and Hyder forced to retreat. The British lost 306 men, while the Mysores are computed to have lost 10,000."

Battle of Sholingur, 1781, First Mysore War

Narrative from Reference 2

"Fought September 27, 1781, between the British, 10,000 strong, under Sir Eyre Coote, and the Mysores, numbering about 80,000, under Hyder Ali. Hyder was surprised in the act of striking camp, and though a series of cavalry charges enabled him to withdraw his guns in safety, it was at a cost of 5,000 men that he eventually made good his retreat. The British loss did not exceed 100."

APPENDIX B

DESCRIPTION OF THE VARIOUS BATTLE GROUPINGS

Battle Grouping	Comprises
Phase-1 Battles	92 battles as listed in Table I of Reference 1.
Phase-2 Battles	83 battles as listed in Table I.
Excluded Battles	2 battles listed as battles number 21 and 22 in Table I.
Modified Battles	81 battles obtained by subtracting the 2 Excluded Battles from the Phase-2 Battles.
Exceptional Battles	15 battles listed as battles number 4, 5, 15, 16, 18, 21, 22, 43, 45, 47, 74, 77, 80, 81, and 82 in Table I.
Censored Battles	68 battles obtained by subtracting the 15 Exceptional Battles from the Phase-2 Battles.
Non-Harbottle Battles	27 battles which did not use Reference 2 as a source and which are listed in Table I as battles number 6, 7, 8, 11, 12, 14, 23, 34, 25, 27, 45, 52, 57, 59, 60, 61, 62, 63, 64, 66, 68, 69, 79, 80, 81, 82, and 83.
Harbottle Battles	56 battles which use Reference 2 as a source, obtained by subtracting the 27 Non-Harbottle Battles from the Phase-2 Battles.

APPENDIX C

SOME NARRATIVE ACCOUNTS OF BATTLES WITH EXTREME VALUES OF RESIDUAL ADVANTAGE

Battle of Bronkhurst Spruit, 1880, First Boer War, Residual Advantage = -1.244

Narrative from Reference 2, Article on "Bronkhurst Spruit"

"The opening engagement of the war, when, on December 20, 1880, a British column, 259 strong, under Colonel Anstruther, was ambushed by 150 Boers under Joubert, and defeated with a loss of 155 killed and wounded. The Boers stated their losses at 2 killed and 5 wounded."

Battle of Atbara, 1898, Soudan Campaigns, Residual Advantage = -0.953

Narrative from Reference 2, Article on "Atbara"

"Fought April 8, 1898, between the British and Egyptian army, 14,000 strong, under Sir Herbert Kitchener, and 18,000 Mahdists under Mahmud. The Mahdists occupied an entrenched zareeba on the Atbara, where they were attacked and utterly routed, with a loss in the zareeba of 5,000 killed and 1,000 prisoners, while many more fell in the pursuit. Mahmud was captured. The Anglo-Egyptian losses were 570 killed and wounded, including 29 British officers."

Narrative from Reference 4

"The forces of the khalifa remaining quiet, the sirdar visited Kassala and negotiated with the willing Italians for its restoration to Egypt. An Egyptian force from Suakin took it over formally on Christmas day 1897. On his return to Berber, the sirdar received information of an intended advance of the khalifa northward. He at once ordered a concentration of Egyptian troops toward Berber, and telegraphed to Cairo for a British

brigade. Disagreement among the khalifa's generals postponed the dervish advance and gave Kitchener much-needed time. But at the end of February, Mahmud crossed the Nile to Shendi with some 12,000 fighting men, and with Osman Digna advanced along the right bank of the Nile to Aliab, where he struck across the desert to Nakheila, on the Atbara. His force consisted of Gatacre's British brigade and Hunter's Egyptian division, with cavalry, a camel corps, and artillery. The dervish army reached Nakheila on March 20, and entrenched themselves. It was ascertained from prisoners that Mahmud's army was short of provisions and Kitchener therefore did not hurry. He sent his flotilla up the Nile and captured Shendi, the dervish depot, on March 27. On April 4 he advanced and, taking the precaution to construct a strong zariba on the night of the 7th, he marched to the attack of Mahmud's zariba, which, after an hour's bombardment in the morning was stormed with complete success. Mahmud was captured with several hundred of his men, and 3,000 were killed. The sirdar lost 80 killed."

Battle of Kwajalein (North), 1944, Second World War,
Residual Advantage = 0.845

Narrative from Reference 6, Volume VII, pp. 230-281, passim

"Two points on Kwajalein Atoll, Roi-Namur in the north and Kwajalein island in the south, were the main objectives of the Marshall Islands operation . . . . .

"Roi-Namur was the enemy's principal air base, Kwajalein Island his principal naval, and Ebeye his seaplane base. The two groups were about 44 miles apart, hence Admiral Turner's plan contemplated two separate but simultaneous amphibious assaults . . . . .

"The Northern Attack Force included plenty of veteran combatant ships, but most of the transports were new, with green crews, and the 4th Marine Division, created since the war began, had never been blooded. The Southern Force, on the contrary, consisted of experienced ships and veteran troops. . . . .

" The Japanese high command had ordered Roi-Namur to be defended to the last man. And the number of defenders there was high for so tiny an area; somewhere between 3,500 and 3,800 men, under naval Captain Seiho Arima . . . . .

" There were several sharp showers during D-day . . . . .

" H-Hour for the landing on Ennuebing and Mellu had been set for 0900 (31 January 1944), but conditions and contretemps delayed the schedule.

" Although the Northern Task Force devoted the greater part of its energy on D-day to securing the five flanking islets, Roi and Namur were the real objective . . . . .

" Three days' bombing and bombardment were not enough to pulverize the Roi-Namur rectangle, roughly 2,300 yards long and 900 wide . . . . .

" It is difficult to see how the defenders could have slept a wink for three nights before the landing. And, in contrast to the 'tough hombres' the Marines had encountered at Tarawa, the defenders of Roi and Namur were pretty well 'pooped out' before any troops landed . . . . .

" At Roi-Namur, if the results (of the bombardment and bombing) were not perfect, a vast amount of damage was done and a large proportion of the 3,700 defenders were killed . . . . . About 6,000 tons of explosives, including aviation rockets, were hurled at Roi and Namur before the assault began, as compared with 2,400 tons at Betio . . . . .

" The four Marine artillery battalions, as we have seen had been laboring all night to get enough pieces and projectiles ashore on the three little islets to permit them to take part in the pounding. Their guns were registered on targets at dawn (of D-day plus 1). Preparatory fires concentrated on the beach for about three hours before the landing, then shifted inland and continued until they began to endanger the rapidly advancing assault troops. These guns fired over 5,000 rounds of 75-mm and 105-mm ammunition on Namur and Roi . . . . .

"At 1157 the first wave landed Marines of the 23rd RCT on the Red beaches at Roi, and about three minutes later the initial echelon of the 24th hit the Green beaches on Namur . . . . .

"Very rapid and vigorous was the action on Roi. The island appeared to be completely deserted and utterly devastated, and the first defenders encountered acted punch-drunk. But about 300 resolute Japanese had survived the terrific bombardment and were hidden in wrecked blockhouses and piles of rubble, mostly along the oceanside . . . . .

"By 1311 the 1st and 2nd Battalions of the 23rd Regiment were in full force on the 0-1 line on Roi, and a regimental command post had been set up on the beach. Between 1530 and 1600 the attack was resumed, so successfully that the north coast was reached promptly, and Colonel L. R. Jones sent a company of riflemen and some medium tanks to Namur, where the going was much tougher. The 23rd RCT then concentrated on mopping up around Roi airfield. Japanese who had taken cover in drainage ditches adjacent to runways fired upward from the ditches and had to be blown out with bangalore torpedoes or satchel charges.

"Thus, enemy opposition at Roi was light, and so were the American casualties . . . . .

"(On Namur) An anti-tank ditch extending behind the beach prevented the amphtracs from proceeding inland as at Roi, so the Marines debarked at the water's edge. Only light resistance developed as they pushed rapidly ahead on foot. Within a quarter of an hour the 24th Marines had pushed 200 yards inland, except for the extreme right where 'Sally Point' was fiercely resisting. Suddenly, about 1245, the crackle of rifle fire and the rat-tat-tat of machine guns were drowned by a tremendous explosion. Fragments of concrete, steel, wood, shrapnel, and even torpedo warheads rained over the surrounding area, killing about 20 Marines and wounding many more . . . . . A blockhouse filled with torpedoes and heavy ammunition had

exploded . . . . . Two less violent explosions occurred on Namur within the next half-hour, and these three accounted for over half the casualties suffered by the 2nd Battalion . . . . .

" At 1820 Colonel Hart ordered his men to make every effort to secure the north shore of Namur before dark . . . . . By 1930 when darkness closed in, the 2nd and 3rd Battalions were about 175 yards north of the 0-1 line. Colonel Hart then ordered his men to establish perimeter defense, hold the ground gained, and prepare to resume attack in the morning . . . . .

" During the darkness small groups of Japanese, some without weapons, infiltrated. Shortly after 0700 February 2, they executed a series of small banzai charges for about half an hour, by the end of which they were liquidated. At 0900 the 3rd Battalion resumed its forward movement with three companies attacking, supported by medium tanks. They quickly overran the remainder of their sector and by noon, 2 February, were resting on the ocean shore of Namur . . . . .

" General Schmidt announced at 1418 February 2 that Namur was secured . . . . .

" The captured Japanese who were interrogated seemed to be imbued with a curious superstition about the Americans. They believed we had a secret weapon . . . . . But the American secret weapon was concentrated and accurate fire power delivered from land, sea, and air; . . . . . It is unlikely that any portion of this globe had ever before received such a concentration of bombs and shellfire. . . . .

At this point Morrison inserts the following footnote:

" Admiral Conolly's Report on Operations in the Northern Sector states that during 8 hours of daylight 31 January his fire support ships fired 23.4 rounds per minute, and during 5 hours of daylight 1 February, 40.9 rounds per minute. This does not include the following ammunition expended by his twelve LCIs: .50 caliber 56,285 rnds; 20-mm 23,700 rnds; 40-mm 13,065 rnds; $4\frac{1}{2}$-inch rockets (smoke) 144 rnds; and $4\frac{1}{2}$-inch rockets (TNT) 2541 rnds."

Battle of Monongahela, 1755, Seven Year's War,
Residual Advantage = -0.812

Narrative from Reference 2, Article on "Monongahela"

"Fought July 9, 1755, between 900 French and Indians, under Contrecour, and about 1,400 British and Virginians, under Braddock. The English were attacked shortly after crossing the river, and though the officers and the Virginians fought gallantly, the troops, ignorant of Indian warfare, gave way to panic, and after three hour's fighting, were driven across the Monongahela, with a loss of 877 killed and wounded. Of 86 officers, 63 fell, including Braddock, who was mortally wounded. The French lost 16 only; their Indian allies somewhat more heavily."

Narrative from Reference 9, p. 7

"Braddock marched his command to Fort Cumberland, where his troops, all told, numbered some 2,200 . . . . . With 1,373 picked men, he moved against Fort Duquesne. On crossing the Monongahela river, about ten miles from the fort, without a single scout to give warning, he fell into an ambuscade set for him by the French and Indians; and, in the dense forest, his command was cut to pieces and routed.

"This happened on the 9th of July 1755. Out of eighty-six British officers, sixty-three were killed or disabled, Braddock being mortally wounded; and only 459 of the rank and file came off unhurt. On the other side, only sixteen white men, and about thirty-five Indians, were killed or wounded."

Battle of Kennesaw Mountain, 1864, American Civil War,
Residual Advantage = 0.762

Narrative from Reference 2, Article on " Kennesaw Mountain"

" Fought June 27, 1864, between 90,000 Federals, under General Sherman, and 50,000 Confederates, under General Johnston. Sherman attacked Johnston in a strong position and was repulsed with a loss of about 3,000, the Confederates losing 500 only."

Narrative from Reference 9, p. 541

" The line now included to crest of Kennesaw Mountain, from end to end."

" On the night of the 21st Johnston, becoming concerned about the pressure on his left, shifted Hood's corps from the right to the left of this line. In this quarter two of Hood's divisions engaged one of Hooker's divisions and one of Schofield's brigade, and suffered a heavy loss on the 22nd. Skirmishing between the hostile lines went on for three or four days without any decisive result. At length, on the 27th of June, Sherman assaulted the strong position on Kennesaw Mountain with the Armies of the Cumberland and the Tennessee, while Schofield threatened Johnston's left. The assault fell mainly upon the corps of Hardee and Loring, and was repulsed with a loss of about 3,000 men. The entire Confederate loss did not exceed 500."

Battle of Wartemberg, 1813, Campaign of Leipsig,
Residual Advantage = 0.827

Narrative from Reference 2, Article on " Wartemberg"

" Fought October 3, 1813, when Blucher, with 60,000 Prussians defeated 16,000 French, under Bertrand, posted in a very strong position, protected by a dyke and a swamp. Aided by the ground, the French

withstood the Prussian attack for over four hours, but finally Blucher turned their right flank and drove them from their position. The Prussians lost about 5,000. The French admit a loss of 500 only."

Battle of Sempach, 1386, War of Sempach
Residual Advantage = 0.901

Narrative from Reference 2, Article on "Sempach"

"Fought July 9, 1386, between 6,000 Austrians, under Duke Leopold, and 1,500 Swiss Confederates. The Swiss gained a complete victory, the Austrians losing 1,500 killed and wounded, while only 120 Swiss fell. The battle is celebrated for the heroic action of Arnold von Winkelreid, who broke the line of the Austrian spearmen at the cost of his life, and enabled his followers to penetrate their phalanx."

Narrative from Reference 3, p. 401

"The important victory of Sempach (1386) finally freed the Swiss from the Austrian House of Hapsburg. Here again the mountaineers were outnumbered, although not so heavily as at Morgartern. At Sempach they had between 1500 and 1600 men against 6,000 Austrians. The terrain was sloping meadowland cut up by hedges and streams. I would remind the reader that open field agriculture had been from the beginning an important factor in the supremacy of cavalry. The Austrians therefore dismounted their men-at-arms and at first shock they drove back the Swiss. Their success, however, was short-lived. We have already seen how impossible it was to advance for any distance on foot while wearing the heavy plate armour of the time, without extreme fatigue. In this case the difficulty of the ground and still more the sun of a hot July day made the task of the cumbersome Austrian men-at-arms still harder. Meantime, the Swiss, unburdened by armour, charged again and again. At last the Hapsburg troops broke and their commander was killed."

Battle of Meeanee, 1843, Scinde Campaign
Residual Advantage = 0.990

Narrative from Reference 2 Article on "Meeanee"

" Fought February 17, 1843, between 2,800 British and native troops under Sir Charles Napier, and about 20,000 Beluchis, under the Amirs of Scinde. The infantry were at one time almost overpowered by the overwhelming numbers of the enemy, who attacked with great bravery, but they were rescued by a charge of the 9th Bengal cavalry, who broke up the assailants, and in the end the Beluchis were routed with a loss of 5,000 men and several guns. The British lost 256 killed and wounded."

REFERENCES

1. Robert L. Helmbold, Historical Data and Lanchester's Theory of Combat, CORG-SP-128, 1 July 1961.

2. Thomas B. Harbottle, Dictionary of Battles, E. P. Dutton & Co., New York, 1905.

3. O. L. Spaulding, H. Nickerson, and J. W. Wright, Warfare: A Study of Military Methods from the Earliest Times, The Infantry Journal, Inc., Washington, D. C., 1937.

4. Encyclopedia Brittanica

5. " Civil War: Part II, Now History, The Battles," LIFE, January 20, 1961.

6. Samuel E. Morison. History of U. S. Naval Operations in World War II, Little, Brown and Co., Boston, 1935.

7. E. L. Crow, F. A. Davis, and M. W. Maxfield, Statistics Manual, Dover Publications, Inc., New York, 1960. (This is a republication of NAVORD Report 3369-NOTS 948.)

8. Furman S. Acton, Analysis of Straight-Line Data, John Wiley & Sons, Inc., New York, 1959.

9. Matthew F. Steele, American Campaigns, Military Service Publishing Co., Harrisburg, Pa., 6th Printing, June, 1949.

10. Robert L. Helmbold, " Lanchester's Equations, Historical Battles, and War Games," included in Proceedings of the Eighth Military Operations Research Symposium; held at Pacific Missile Range, Point Mugu, California, October 18-21, 1961.

11. R. P. Mulholland, and R. D. Specht, The Rate of Advance of the Front Line in Some World War II Campaigns, RAND-RM-1072, 16 April 1953.